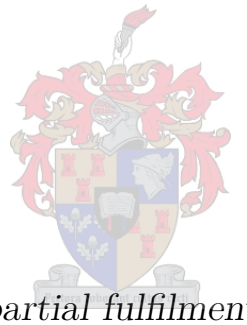
Explicit bound on Siegel zeros of imaginary quadratic fields

by

Faratiana Brice Razakarinoro



Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics in the Faculty of Science at Stellenbosch University

Supervisor: Dr. Dimbinaina Ralaivaosaona

December 2019

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2019

Abstract

Explicit bound on Siegel zeros of imaginary quadratic fields

F. B. Razakarinoro

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MSc

December 2019

Many problems in number theory depend on the zeros of Dirichlet L-functions. The Generalized Riemann Hypothesis (GRH) states that the zeros of the L-functions with positive real parts lie precisely on the vertical line $\mathrm{Re}(s) = \frac{1}{2}$. Like the classical Riemann Hypothesis, the GRH has profound consequences about the distribution of prime numbers. It is known that the L-functions associated with characters defined by the Kronecker's symbol $\chi_\Delta(n) = (\frac{\Delta}{n})$ might have zeros in the intervals $(1 - \frac{c}{\log|\Delta|}, 1)$ where $c > 0$ is an absolute constant. These hypothetical zeros are known as Siegel zeros or sometimes called Landau-Siegel zeros. In this thesis, we consider the character χ_Δ where $\Delta < 0$. The study of these characters is naturally related to the study of quadratic forms and imaginary quadratic fields.

In 1975, Goldfeld and Schinzel showed that there exists a positive constant $\mathcal{C}$ such that if $L(\beta, \chi_\Delta) = 0$ and $\beta > 0$, then $\beta < 1 - \frac{\mathcal{C}}{\sqrt{|\Delta|}}$. In this thesis, we use recent computational results, including Watkins' classification of negative fundamental discriminants with class number up to 100, to make the constant $\mathcal{C}$ in Goldfeld-Schinzel's bound explicit.

Uittreksel

Eksplisiete grens op Siegel nul van imaginêre kwadratiese liggame

("Explicit bound on Siegel zeros of imaginary quadratic fields")

F. B. Razakarinoro

Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MSc

Desember 2019

Baie probleme in getalleteorie hang van die nulpunte van Dirichlet L-funksies af. Die Veralgemeende Riemann-hipotese (VRH) staat dat die die nulpunte van die L-funksies met positiewe reële gedeeltes presies op die vertikale lyn $\mathrm{Re}(s) = \frac{1}{2}$ lê. Soos die klassieke Riemann-hipotese, het die VRH diep gevolge oor die verspreiding van die priemgetalle. Dit is bekend dat die L-funksies geassosieer met karakters wat gedefinieer w ord d eur d ie K ronecker simbool $\chi_\Delta(n) = (\frac{\Delta}{n})$ wel nulpunte in die intervalle $(1 - \frac{c}{\log|\Delta|},\ 1)$, waar $c > 0$ 'n absolute konstante is, kan hê. Hierdie hipotetiese nulpunte is bekend as Siegel nulpunte of somtyds word daarna verwys as Landua-Siegel nulpunte. In hierdie tesis beskou ons die karakter χ_Δ, waar $\Delta < 0$. Die studie van hierdie karakters hou natuurlik verband met die studie van kwadratiese vorme en imaginêre kwadratiese liggame.

In 1975 het Goldfeld en Schinzel bewys dat daar 'n positiewe konstante $\mathcal{C}$ bestaan sodanig dat, indien $L(\beta, \chi_\Delta) = 0$ en $\beta > 0$, dan is $\beta < 1 - \frac{\mathcal{C}}{\sqrt{|\Delta|}}$. In hierdie tesis gebruik ons onlangse resultate, insluitende Watkins se klassifikasie van negatiewe fundamentele diskriminante met klasgetal tot 100, om die konstante $\mathcal{C}$ in Goldfeld-Schinzel se grens eksplisiet te maak.

Acknowledgements

First and foremost, I would like to thank God Almighty for giving me the time, health and knowledge to undertake this research study. Without his blessings, this achievement would not have been possible.

I would like to express my sincere gratitude to my supervisor Dr. Dimbinaina Ralaivaosaona: thank you for your guidance, encouragement and close collaboration throughout this project. Your mathematical experiences really helped and taught me innumerable lessons. Thank you also for your advice for improving the writing of this manuscript.

My deep thanks also go to the ALGANT Consortium for the financial support, the University of Padova where I spent a year, and the Stellenbosch University where I completed my last year of the program.

I am very thankful to Prof. Florian Breuer for helping and supporting me to be part of the program and also for suggesting me the research project.

I am also very grateful to all the persons who helped me and made my stay pleasant while I was in Italy as well as in South Africa. Special thanks to Dr. R. Benjamin for translating the abstract into Afrikaans.

Last, but certainly the most important, I would like to thank my entire family, especially Mirana, for providing me continuous encouragement and support. Thank you all!

Brice

Dedication

To Mirana

Contents

Notations

$\mathbb{Z}$	the ring of rational integers
$\mathbb{Q}, \mathbb{C}$	the field of rational numbers and complex numbers
$N(\alpha)$, $N(\mathfrak{a})$	the norm of an element α and an ideal $\mathfrak{a}$
χ_0	the principal character
$\omega (= \omega_K)$	the number of automorphs of a quadratic form
$r_{[Q]}(n)$	the representation number of n by the quadratic form Q
$\mathcal{C}(d)$	the form class group for a given discriminant d
$h(d)$	the order of $\mathcal{C}(d)$, i.e. the class number for the discriminant d
O_K	the ring of integers of the number field K
$\mathcal{I}_K$	the group of fractional ideals of O_K
$\mathcal{P}_K$	the subgroup of principal fractional ideals of O_K
$\mathcal{C}(O_K)$	the ideal class group for the number field K
$h(O_K)$	the order of $\mathcal{C}(O_K)$, i.e. the class number for the ring O_K
$(\frac{d}{p})$	the Legendre symbol (for an odd prime p)
χ_d	the Kronecker symbol which is also a primitive character modulo $\lvert d \rvert$
$L(s, \chi)$	the Dirichlet L-function associated to the character χ
ζ_K	the Dedekind zeta function associated to the number field K
$\zeta(s, F)$	the Epstein zeta function for the quadratic form $F(x, y)$
β	the exceptional zero of the Dirichlet L-function, i.e. the Siegel zero
$\Gamma(s)$	the Gamma function
$\varphi(q)$	the Euler-totient function of q
$E_{/K}$	the elliptic curve E over the field K
$E(K)$	the group of K-rational points of the elliptic curve $E_{/K}$
$E(K)[tors]$	the torsion subgroup of the elliptic curve $E_{/K}$
$\#E(K)[tors]$	the size of the torsion subgroup of the elliptic curve $E_{/K}$
$E[m]$	the m-torsion subgroup of the elliptic curve $E_{/K}$
$[n]$	the multiplication-by-n map
$\mathrm{End}_K(E)$	the ring of endomorphisms of $E(K)$
$[L : K]$	the degree of the field extension L/K
$\mathrm{Gal}(L/K)$	the Galois group of the extension L/K
$T_{CM}(d)$	the largest $\#E(K)[tors]$ where $E_{/K}$ is of CM-type and $d = [K : \mathbb{Q}]$
$\mathcal{I}_K(\mathfrak{m})$	the group of fractional ideals of O_K coprime to the modulus $\mathfrak{m}$
$\mathcal{P}_{K,1}(\mathfrak{m})$	the subgroup of $\mathcal{I}_K(\mathfrak{m})$ generated by some specific principal ideals of O_K
$\left(\frac{L/K}{\mathfrak{P}}\right)$	the Artin symbol of the prime $\mathfrak{P}$ of L
$\Phi_{L/K,\mathfrak{m}}$	the Artin map for L/K and $\mathfrak{m}$
$K^{(\mathfrak{a})}$	the ray class field for $\mathfrak{a}$
$\nu(a)$	the number of representations of a as $N(\mathfrak{a})$ for an ideal $\mathfrak{a}$
$\omega(n)$	the number of distinct prime divisors of n
$\tau(\chi)$	the Gauss sums associated with the character χ

Chapter 1

Introduction

1.1 History

One of the oldest problem in number theory is *how to write integers as sums of squares*. Many representation theorems of integers were proved in the eighteenth century. In 1654, Fermat stated the following: *let p be a prime number*,

$$\text{if } p = 6n + 1 \text{ then } p = x^2 + 3y^2,$$
$$\text{if } p = 8n + 1 \text{ then } p = x^2 + 2y^2.$$

These were proved by Euler in 1761 and 1763. Nine years later, Euler also observed that *for $x = 1, 2, \ldots, 40$, we have*

$$x^2 - x + 41 = p,$$

where p is a prime number. To handle the general problem, it was in 1773 that Lagrange introduced, for the first time, the theory of binary quadratic forms, that is the question of when an integer n can be written in the form

$$n = ax^2 + bxy + cy^2,$$

where a, b, c are integers. The quantity $b^2 - 4ac$ is defined to be the discriminant of the quadratic form $ax^2 + bxy + cy^2$. Lagrange also developed the concept of equivalence of forms as well as the notion of reduction for binary quadratic forms. This idea was further explored by Gauss. In 1801, with the publication of Gauss's Disquisitiones Arithmeticae, Gauss defines the composition of quadratic forms and proves that the classes of binary quadratic forms with a given discriminant has a group structure with composition as a group law. Gauss even conjectured that *the number of negative discriminants which have a given class number is finite.* This is the famous *class number problem* or the *Gauss class number problem for imaginary quadratic fieds*. In its modern version, the statement of the class number problem consists of giving for each

$n \geq 1$ a complete list of imaginary quadratic fields having class number n. Initially, Gauss stated that *the class number $h(\Delta)$ of an imaginary quadratic field of discriminant $-\Delta$ tends to infinity with Δ*. A complete proof of this was given by Heilbronn in 1934 after extending a work of Deuring (1933) and Mordell (1934). Then, the case of the *low class number lists* were showed by Baker, Stark, Heegner and Oesterlé. By this term, we mean class number $h(\Delta)$ equals to 1, 2, 3. Moreover, a theorem of Siegel implies that for any $\varepsilon > 0$, there exists a constant $k(\varepsilon)$ such that

$$h(\Delta) > k(\varepsilon)|\Delta|^{\frac{1}{2}-\varepsilon},$$

for Δ negative. The constant in this bound is not effective. Finally in 1980, Goldfeld, Gross and Zagier obtained an unconditional estimate with an effective constant on the class number. In summary, Goldfeld found a lower bound on the class number involving the existence of an elliptic curve with some conditions. On the other side, Gross and Zagier proved in their work the existence of the elliptic curve considered by Goldfeld. Many years later, it was in 2004 that the case $h(\Delta)$ up to 100 was proved by Watkins in his paper [27]. This is the last known result so far on the class number problem.

On the other hand, a remarkable fact about the class number is that it is known to have a close relation with the Siegel zero. For example, this can be seen in the paper [12] of Goldfeld in which he provides an asymptotic formula relating the Siegel zero and the class number. Moreover, we also have a result of Hecke which says that: *Let χ be a real primitive character modulo $|\Delta|$. Suppose there exists $\mathcal{K}$ such that no $L(s,\chi)$ has a zero β satisfying*

$$1 - \frac{\mathcal{K}}{\log|\Delta|} < \beta < 1.$$

Then there is some other constant $\hat{\mathcal{K}}$ such that

$$h(\Delta) > \frac{\hat{\mathcal{K}}\sqrt{|\Delta|}}{\log|\Delta|}.$$

As one can see, Hecke's Theorem gives us a concrete statement showing the dependence between the class number $h(\Delta)$ and the Siegel zero β. In addition to this, by the Dirichlet's Class Number Formula, the situation becomes more interesting as we may know some informations on $L(1,\chi)$, then on $h(\Delta)$ as well. In other words, finding a bound of $L(1,\chi)$ is essentially the same concept as localizing the Siegel zero β.

The concept of Siegel zero comes from the following theorem.

Theorem 1.1.1. *[9, p.93] Let $\mathcal{R}$ be the region defined by*

$$\mathcal{R} = \left\{ s = \sigma + it \in \mathbb{C} \mid \sigma \geq 1 - \frac{C}{\log \Delta|t|} \text{ if } |t| \geq 1 \text{ and } \sigma \geq 1 - \frac{C}{\log \Delta} \text{ if } |t| \leq 1 \right\}.$$

Then there exists a positive absolute constant C, with the following property:

- *If χ is a complex character modulo Δ, then $L(s,\chi)$ has no zero in the region defined by $\mathcal{R}$,*
- *If χ is a real non-principal character, the only possible zero of $L(s,\chi)$ in the region $\mathcal{R}$ is a single (simple) real zero β.*

More precisely, we are concerned with the second assertion of this theorem. Hence,

Definition 1.1.2. We call *Siegel zero* of $L(s,\chi)$ the exceptional zero β defined as in Theorem 1.1.1.

Recently, a new version of the previous theorem has been given by H. Kadiri in [18, Theorem 1.1], which states that: *Let Δ be an integer with $3 \leqslant \Delta \leqslant 400\,000$ and χ a non-principal primitive character modulo Δ. Then the Dirichlet L-function $L(s,\chi)$ does not vanish in the region*

$$\mathrm{Re}(s) \geqslant 1 - \frac{1}{5.60 \log\left(\Delta \max\{1, |\mathrm{Im}(s)|\}\right)}.$$

In particular, we can observe that we are most concerned with the real primitive Dirichlet character. The reason for this is that in order to study Siegel zeros, it suffices to consider primitive characters. This can be seen from the following result in [1, Theorem 12.9]: *Let χ be any Dirichlet character modulo Δ. Then for all s we have*

$$L(s,\chi) = L(s,\psi) \prod_{p|\Delta} \left[1 - \frac{\psi(p)}{p^s}\right],$$

where ψ is a primitive character.

1.2 Known results

From Theorem 1.1.1, we have that Siegel zero is on the real axis. In fact, it is known to be close to 1. In 1935, Page showed the following:

Theorem 1.2.1. *If χ is a real non-principal character modulo Δ, and if β_1 and β_2 are real zeros of $L(s,\chi)$, then there is a positive constant c_1 such that*

$$\min\,(\beta_1, \beta_2) \leqslant 1 - \frac{c_1}{\log \Delta}.$$

As for the lower bound of the distance of β from 1, in 1975, Goldfeld and Schinzel proved in [13] that:

Theorem 1.2.2. *For any fundamental discriminant $\Delta < 0$, we have*

$$1 - \beta \geq \frac{C}{\sqrt{|\Delta|}}. \tag{1.2.1}$$

And as $\Delta \to -\infty$, we have

$$1 - \beta \geq \left(\frac{6}{\pi} + o(1)\right)\frac{1}{\sqrt{|\Delta|}}.$$

A similar result was also given by Pintz [21] and Haneke [14]. About recent computational results, Watkins showed in his paper [28] of 2003 that: *Let χ be a real odd Dirichlet character of modulus $|\Delta| \leq 300\,000\,000$. Then $L(s, \chi)$ has no positive real zeros.* This result is indeed an extension of the work of Low (1969) and Purdy (1972) which says that *if $|\Delta| \leq 800\,000$ and $|\Delta| \neq 115\,147, 357\,819, 636\,184$ then $L(s, \chi)$ has no positive real zeros.* More recently, on November 2018, M.A. Bennett, G. Martin, K. O'Bryant and A. Rechnitzer showed in [11] that:

Proposition 1.2.3. *Let $\Delta \geq 3$ be an integer, and let χ be a quadratic character modulo Δ. If $\beta > 0$ is a real number for which $L(\beta, \chi) = 0$, then*

$$\beta \leq 1 - \frac{40}{\sqrt{\Delta}\log^2\Delta}.$$

Then some days later, a new result of T. Morrill and T. Trudgian which appears in [20] states that:

Theorem 1.2.4. *If χ is a real non-principal character modulo Δ, with $\Delta \geq 3$, and if β_1 and β_2 are real zeros of $L(s, \chi)$, then we have*

$$\min\,(\beta_1, \beta_2) \leqslant 1 - \frac{1.011}{\log\,\Delta}.$$

In this work, our goal is to provide an explicit version of Equation (1.2.1). We show that

Theorem 1.2.5. *Let $\Delta < 0$ be a fundamental discriminant and let $L(s, \chi_\Delta)$ be the L-function associated with the primitive character $\chi_\Delta(n) = (\frac{\Delta}{n})$. If there is $\beta > 0$ such that $L(\beta, \chi_\Delta) = 0$ then*

$$1 - \beta \geq \frac{1.151}{\sqrt{|\Delta|}}.$$

This thesis is organized as follows. In Chapter 2, we develop some results concerning imaginary quadratic fields. We present the basics and properties of quadratic forms, quadratic fields and primitive characters. We also show

the Dirichlet Class Number Formula for negative fundamental discriminant. In Chapter 3, we present a result due to Clark and Pollack concerning the size of the torsion subgroup of a CM elliptic curve. Their approach relies on a bound of the form as in Equation (1.2.1). The proof of Theorem 1.2.5 is done in Chapter 4.

Chapter 2

Imaginary Quadratic Fields

Quadratic fields are extension of $\mathbb{Q}$ of degree 2. An imaginary quadratic field is described as the field $\mathbb{Q}(\sqrt{d})$ where d is a negative integer. Quadratic fields have been studied in great depth, initially as part of the theory of binary quadratic forms. More specifically, one of the most famous problem concerning imaginary quadratic fields is the *Gauss class number problem for imaginary quadratic fieds*. On the other hand, there is also the well-known *class number formula* which is attribuated to Dirichlet. It was first conjectured by Jacobi in 1832 and showed in full by Dirichlet in 1839. The importance of this formula lies on the fact that it relates some invariants of the quadratic fields with the value of the L-functions at $s = 1$.

In this chapter, we begin with the study of quadratic forms. We then introduce the concept of primitive characters which intervene in the Dirichlet class number formula. Also, we are going to explore the quadratic fields, in particular its ring of integers.

2.1 Quadratic Forms

A homogeneous polynomial is a polynomial whose nonzero terms all have the same degree. Quadratic forms are a special case of such general concept.

The study of quadratic forms dates back many centuries. In particular, the question of whether a given integer can be represented by a quadratic form is one of the oldest problems in number theory.

2.1.1 Equivalence class of forms

Definition 2.1.1. An *integral binary quadratic form* is an expression of the form

$$Q(x, y) = ax^2 + bxy + cy^2,$$

where $a, b, c \in \mathbb{Z}$. For convenience, we will use the term *quadratic form* or sometimes *form* to be short. We define the *discriminant* of $Q(x, y)$ as $b^2 - 4ac$ and usually denote it by d.

Definition 2.1.2. Let $Q(x, y)$ be a quadratic form. We say that $Q(x, y)$ is a *primitive form* if the coefficients a, b, c are relatively prime.

In some case, the quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ is represented by $[a, b, c]$ or in matrix notation

$$\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}. \tag{2.1.1}$$

In the matrix representation, the form $Q(x, y)$ is obtained as follows

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + bxy + cy^2. \tag{2.1.2}$$

Also, note that the discriminant d of $Q(x, y)$ is given by

$$-4 \det \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}. \tag{2.1.3}$$

With some calculation on $Q(x, y)$, we find that

$$4a(ax^2 + bxy + cy^2) = (2ax + by)^2 - dy^2.$$

If we have $a \neq 0$ then we may write

$$Q(x, y) = \frac{1}{4a}\Big((2ax + by)^2 - dy^2\Big), \tag{2.1.4}$$

and similarly if $c \neq 0$, then we obtain

$$Q(x, y) = \frac{1}{4c}\Big((2cy + bx)^2 - dx^2\Big).$$

A simple remark from the above calculation is that $Q(x, y)$ can take positive and negative values for any real x and y if $d > 0$. Also, in the case $d < 0$, for any real x and y we have $Q(x, y) > 0$ (resp. $Q(x, y) < 0$) when $a > 0$ (resp. $a < 0$) by using Equation (2.1.4).

Definition 2.1.3. Let d be the discriminant of the form $Q(x, y)$.

- When $d > 0$, the form $Q(x, y)$ is called *indefinite form*.
- When $d < 0$, the form $Q(x, y)$ is called *positive definite form* if $a > 0$ and *negative definite form* if $a < 0$.

Now, we recall the following:

Definition 2.1.4. An interger d is said to be a *fundamental discriminant* if either d is squarefree and $d \equiv 1 \bmod 4$, or $d = 4n$ for some squarefree integer $n \equiv 2$ or $3 \bmod 4$.

In particular, we have

Proposition 2.1.5. *Every quadratic form $Q(x, y)$ with a fundamental discriminant d is primitive.*

Proof. Write $Q(x, y) = ax^2 + bxy + cy^2$ and $d = b^2 - 4ac$. Assume that $\gcd(a, b, c) = k > 1$. Then, we have $d = (kb')^2 - 4(ka')(kc') = k^2(b'^2 - 4a'c')$ where $a = ka'$, $b = kb'$ and $c = kc'$, that is, d is not squarefree. Therefore, if d is a fundamental discriminant then $Q(x, y)$ is primitive. □

One of the most classical problem in number theory is the representation of a given integer by a quadratic form. This was the first problem concerning binary quadratic forms.

Definition 2.1.6. We say that an integer m is *represented* by a form $Q(x, y)$ if the equation $Q(x, y) = m$ has integers solutions in x and y. When we have $\gcd(x, y) = 1$, we say that m is *properly represented* by the form $Q(x, y)$.

Based on the above definition, there is a natural equivalence relation in the set of quadratic forms.

Definition 2.1.7. Two quadratic forms $Q(x, y)$ and $P(x, y)$ are said to be *equivalent* if there exist intergers p, q, r and s such that we have

$$Q(x, y) = P(px + ry, qx + sy) \qquad \text{and} \qquad ps - qr = \pm 1. \tag{2.1.5}$$

Clearly, if two forms are equivalent then they represent the same set of integers.

Remark 1. In the above definition, the equality $Q(x, y) = P(px + ry, qx + sy)$ can also be viewed as

$$Q(x, y) = P\Big(\begin{bmatrix} x & y \end{bmatrix} M \Big) \quad \text{where} \quad M = \begin{bmatrix} p & q \\ r & s \end{bmatrix}. \tag{2.1.6}$$

In addition, note that the second relation in Equation (2.1.5) says that the matrix M belongs to the group of invertible integers matrices $\mathrm{GL}(2, \mathbb{Z})$. That is, it makes sense to consider the equivalence of forms as an equivalence relation. In fact, we can say more by considering the subgroup $\mathrm{SL}(2, \mathbb{Z})$ which is known as the modular group.

Definition 2.1.8. Let $Q(x, y)$ and $P(x, y)$ be quadratic forms satisfying Equation (2.1.5). We say that $Q(x, y)$ and $P(x, y)$ are:

- *properly equivalent* when $ps - qr = 1$,
- *improperly equivalent* when $ps - qr = -1$.

In other words, we may also consider a proper equivalence of forms as an equivalence relation since in this case the matrix $\left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$ belongs to $\mathrm{SL}(2, \mathbb{Z})$. In fact, the notion of proper equivalence leads us to the class of forms. We say that *two forms lie in the same class* if they are properly equivalent.

A result which relates the notion of proper equivalence and proper representation is the following.

Lemma 2.1.9. *An integer m is properly represented by a form $Q(x, y)$ if and only if $Q(x, y)$ is properly equivalent to the form $mx^2 + b'xy + c'y^2$ for some $b', c' \in \mathbb{Z}$.*

Proof. Suppose that $m = Q(u, v)$ with relatively prime u and v. We may find integers r, s such that $ur - vs = 1$. Now, write $Q(x, y) = ax^2 + bxy + cy^2$. Then, we have

$$\begin{aligned} Q(ux + ry, vx + sy) &= a(ux + ry)^2 + b(ux + ry)(vx + sy) + c(vx + sy)^2, \\ &= Q(u, v)x^2 + (2aur + bus + brv + 2cvs)xy + Q(r, s)y^2, \\ &= mx^2 + b'xy + c'y^2. \end{aligned}$$

Conversely, plugging $(x, y) = (1, 0)$ we find that m is properly represented by $mx^2 + b'xy + c'y^2$, and so by $Q(x, y)$. □

In terms of discriminant, we can easily prove that two equivalent forms must have the same discriminant. Indeed, given two equivalent forms $Q(x, y)$ and $P(x, y)$ with discriminant d and d' respectively, we have that, for integers p, q, r, s,

$$\begin{aligned} Q(x, y) &= P\Big(\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} \Big), \\ &= \begin{bmatrix} x & y \end{bmatrix} M \begin{bmatrix} a' & b'/2 \\ b'/2 & c' \end{bmatrix} {}^{tr}M \begin{bmatrix} x \\ y \end{bmatrix}, \end{aligned}$$

and the last equality is obtained by writing $P(x, y)$ as in (2.1.2).

In other words, the form $P(x, y)$ can be represented, in matrix form, as $M\left[\begin{smallmatrix} a' & b'/2 \\ b'/2 & c' \end{smallmatrix}\right] {}^{tr}M$, which has a discriminant

$$-4 \det(M)^2 \det \begin{bmatrix} a' & b'/2 \\ b'/2 & c' \end{bmatrix},$$

by using (2.1.3). Recall that $\det(M) = ps - qr$, so we have that

$$d = (ps - qr)^2 d',$$

by looking at the discriminant of the forms. In other words, we have proved that properly equivalent forms have the same discriminant.

2.1.2 Reduced forms

We saw previously that quadratic forms can be splitted into classes of forms. Here, we are interested in the representatives of the classes.

Lemma 2.1.10. *Every quadratic form is equivalent to a form $ax^2 + bxy + cy^2$ with*

$$|b| \leq |a| \leq |c|. \tag{2.1.7}$$

Proof. Let $Q(x, y) = ax^2 + bxy + cy^2$ be a form which is properly equivalent to the given one such that $|b|$ is chosen to be small as possible. For any m, we have

$$\begin{aligned} P(x, y) &= Q(x + my, y) \\ &= a(x + my)^2 + b(x + my)y + cy^2 \\ &= ax^2 + (2am + b)xy + (am^2 + bm + c)y^2 \\ &= ax^2 + (2am + b)xy + c'y^2, \end{aligned}$$

which is properly equivalent to $Q(x, y)$.Assume by contradiction that $|a| < |b|$. Then, there exists m such that $|2am + b| < |a|$. Hence, this contradicts the minimality of $|b|$, and so $|a| \geq |b|$. By a similar argument, we obtain that $|c| \geq |b|$. Now, if $|a| > |c|$, the matrix of proper equivalence $\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ interchanges the coefficients a and c. Therefore, the resulting form satisfies the inequality. □

We note that the above lemma is due to Lagrange. Moreover, it plays an important role in showing the next result.

Proposition 2.1.11. *Let d be a squarefree integer. There are only a finite number of equivalence classes of quadratic forms of discriminant d.*

Proof. By Lemma 2.1.10, we may choose from each equivalence class some quadratic form $ax^2 + bxy + cy^2$ such that $|b| \leq |a| \leq |c|$. Then the total number of equivalence classes is less or equal to the size of the set

$$S := \{ax^2 + bxy + cy^2 \mid d = b^2 - 4ac,\ |b| \leq |a| \leq |c|\}.$$

For every quadratic form in the set S, we have $4a^2 \leq 4|ac| = |d - b^2| \leq |d| + a^2$, hence $|a| \leq \sqrt{\frac{|d|}{3}}$. Moreover, since $|b| \leq |a|$ it follows that there are only finitely many choices of a, b and for any such choice there is at most one integer c such that $d = b^2 - 4ac$. Hence the cardinality of the set S is finite. □

Before we describe the class of quadratic forms, we recall the following.

Lemma 2.1.12. *Let a be a positive integer. Assume that the quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ satisfies (2.1.7). Then, a is the minimum of $Q(x,y)$, meaning that*

$$Q(x,y) \geq a \quad \text{for all } (x,y) \neq (0,0). \tag{2.1.8}$$

Furthermore, ac is the minimum of products of values of Q, meaning that, for all pairs of lattice points (x,y) and (u,v), which are not colinear, one has

$$Q(x,y)Q(u,v) \geq ac. \tag{2.1.9}$$

Proof. First, note that a is represented by $Q(x,y)$ as $a = Q(1,0)$. In the same way, $ac = Q(1,0)Q(0,1)$. Suppose $x \neq 0$, then $Q(x,0) = ax^2 \geq a$. Now, assuming $y \neq 0$, we have $Q(0,y) = cy^2 \geq c$. More generally, if x and y are not 0, we see that

$$\begin{aligned} Q(x,y) &= ax^2 + bxy + cy^2 \\ &\geq ax^2 - |b||x||y| + cy^2 \\ &\geq ax^2 - |b||x||y| + cy^2 - a(|x| - |y|)^2 \\ &= (2a - |b|)|x||y| + (c-a)y^2 \\ &\geq (2a - |b|) + (c-a) \\ &= a + c - |b| \geq c, \qquad \text{since } a \geq |b|. \end{aligned}$$

Assume now that (x,y) and (u,v) are not colinear. Thus, by doing the same procedure as above, we have that the product $Q(x,y)Q(u,v) \geq ac$. □

The next statement provides us a precise description of the representatives of the class of quadratic forms.

Theorem 2.1.13. *Every quadratic form of discriminant d is equivalent to exactly one form $ax^2 + bxy + cy^2$ which satisfies*

$$|b| \leq a \leq c, \text{ and } b \geq 0 \text{ if either } |b| = a \text{ or } a = c. \tag{2.1.10}$$

Proof. When $b = 0$, then it is clear that $ax^2 + bxy + cy^2$ and $ax^2 - bxy + cy^2$ are equivalent. In the case $a = c$, the quadratic form $ax^2 + bxy + cy^2$ is equivalent to $cx^2 - bxy + ay^2 = ax^2 + bxy + cy^2$ via the matrix $\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$. Next, if $|b| = a$ then the equivalence of $ax^2 + bxy + cy^2$ and $ax^2 - bxy + cy^2$ is obtained via the matrix $\left[\begin{smallmatrix} 1 & 0 \\ -sign(b) & 1 \end{smallmatrix}\right]$, with $sign(b) = \pm 1$. Considering Lemma 2.1.10, we can say that every form is equivalent to at least one form satifying the condition (2.1.7).

Now, let us show that those forms satisfying the condition (2.1.7) lie in distinct class of forms, and then we are done. Assume that we have two equivalent forms $Q(x, y)$ and $P(x, y)$ both satisfying (2.1.7). Write

$$Q(x, y) = ax^2 + bxy + cy^2$$
$$P(x, y) = a'x^2 + b'xy + c'y^2.$$

Using Lemma 2.1.12, it follows that a and a' are respectively the minimum of $Q(x, y)$ and $P(x, y)$, that is $a = a'$. In the same way, the minimum of products of pairs are equal, that $c = c'$ as we have $ac = a'c'$. Using the fact that they have the same discriminant, we deduce that $b' = \pm b$. So now we got two cases to treat. First, when $b' = -b$ then this is as we desire so we are done. Secondly, when $b' = b$, it means that the sign of b' and b are opposite. Without loss of generality, say b has negative sign. Then, it follows that $0 < |b| < a < c$. Since $Q(x, y)$ and $P(x, y)$ are equivalent, there exist integers p, q, r, s such that

$$Q(x, y) = P(px + ry, qx + sy).$$

So, we have $a = Q(1, 0) = P(p, q)$ and then we deduce that $q = 0$ and $p = \pm 1$. In the same way, $c = Q(0, 1) = P(r, s)$ implies $r = 0$ and $s = \pm 1$. Finally, in order to have a matrix of determinant 1, the matrix has to be $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ or $\left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$. In other words, $Q(x, y)$ and $P(x, y)$ are the same form, not only equivalent as we have supposed ealier. □

Now, we are mainly concerned with those forms satisfying Condition (2.1.10). In fact, this allows us to classify forms.

Definition 2.1.14. A quadratic form $Q(x, y)$ is called *reduced* if it satisfies Condition (2.1.10).

According to Theorem 2.1.13, one can deduce that each class of quadratic forms contains precisely one reduced form. On the other hand, Proposition 2.1.11 tells us that there are only a finite number of classes. Hence, we can say that there are only a finite reduced forms of given discriminant.

Definition 2.1.15. Given a fundamental discriminant d, the *class number* is the number of equivalence classes of primitive quadratic forms which are positive definite or indefinite. It is usually denoted by $h(d)$, where d is the discriminant of quadratic forms.

Remark 2. We saw that if the discriminant d is negative, then the form may take positive or negative values depending on the sign of a. In fact, the negative definite form is obtained by taking the negation of the coefficient of the positive definite form. That is, we may only consider the positive definite which means those form having $a > 0$. On the other hand, the form is indefinite when $d > 0$. Recall by Lemma 2.1.9 that we may choose some positive integer a, which is

properly represented by the form and also the first coefficient of some properly equivalent form. Thus, it is reasonable to take a representative of each class of forms with the condition that $a > 0$. Hence, it is convenient to define the class number as in Definition 2.1.15 .

An important fact about the class number is that it is at least 1, that is, we have

$$h(d) \geq 1. \tag{2.1.11}$$

This is because there is always at least one quadratic form of discriminant d defined by

$$x^2 - \frac{d}{4}y^2 \quad \text{if } d \equiv 0 \bmod 4, \tag{2.1.12}$$

$$x^2 + xy - \frac{1-d}{4}y^2 \quad \text{if } d \equiv 1 \bmod 4. \tag{2.1.13}$$

These forms are known as the *principal form.*

Definition 2.1.16. We define an *automorph* of a quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ to be the unimodular substitution $M \in \mathrm{SL}(2, \mathbb{Z})$ with

$$\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} = M \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} {}^{tr}M.$$

Proposition 2.1.17. *[24, p.305] The set of automorphs of a primitive quadratic form $Q(x, y)$ of discriminant d is given by*

$$\left\{ \begin{bmatrix} (t-bu)/2 & -cu \\ au & (t+bu)/2 \end{bmatrix} \mid t, u \in \mathbb{Z} \text{ and } t^2 - du^2 = 4 \right\} \subset SL(2, \mathbb{Z}). \tag{2.1.14}$$

The latter proposition allows us to enumerate the automorphs of a quadratic forms with negative discriminant.

Proposition 2.1.18. *Every quadratic form of discriminant $d < 0$ has exactly ω automorphs given by:*

$$\omega = \begin{cases} 2 & \text{if } d < -4, \\ 4 & \text{if } d = -4, \\ 6 & \text{if } d = -3. \end{cases} \tag{2.1.15}$$

Proof. Let $Q(x, y)$ be a quadratic form with a negative discriminant d. The proof is based on the Proposition 2.1.17. If $d < -4$, we have exactly two integer solutions $(\pm 2, 0)$ for the equation $t^2 - du^2 = 4$. As we have $d \equiv 0, 1 \pmod 4$ we get the case $d = -3$ and $d = -4$. When $d = -4$, there are exactly four integer solutions for the equation $t^2 - du^2 = 4$, namely $(\pm 2, 0)$ and $(0, \pm 1)$. Finally, when $d = -3$, the equation $t^2 - du^2 = 4$ has exactly six integer solutions which are $(\pm 2, 0)$ and $(\pm 1, \pm 1)$. □

Definition 2.1.19. Let $Q(x,y)$ be a quadratic form with negative discriminant and let n be an integer. We define the *representation number* of n by the form Q as

$$r_{[Q]}(n) := \frac{1}{\omega}\#\{(x,y) \text{ relatively prime such that } Q(x,y) = n\}, \tag{2.1.16}$$

where ω is defined by (2.1.15) and $[Q]$ denotes the equivalence class containing Q.

Note that $r_{[Q]}(n)$ depends only on the class of Q. Hence, a more general definition is as follows

$$r_d(n) := \sum_{[Q]} r_{[Q]}(n), \tag{2.1.17}$$

where d is a negative discriminant.

The following result on $r_d(n)$ can be found in [24, p.306].

Theorem 2.1.20. *For a negative fundamental discriminant d, we have*

$$r_d(n) = \begin{cases} 0 & \text{if } p^2 \mid n \text{ for some } p \mid d, \\ \prod\limits_{\substack{p\mid n \\ (p,d)=1}} \left(1 + \left(\frac{d}{p}\right)\right) & \text{otherwise.} \end{cases} \tag{2.1.18}$$

2.1.3 Form class group

Here, we present a brief discussion on the composition of quadratic forms. This is mainly due to the work of Gauss by introducing the notion of *direct composition* (see[8]).

Gauss' theory of composition is that for a fixed discriminant, direct composition makes the set of classes of forms into a finite abelian group. But since the notion of direct composition is not easy to work with, most of standard textbooks on quadratic forms make use of different approach. Here, we follow Dirichlet's study of composition. Before giving Dirichlet's definition, we need the following lemma:

Lemma 2.1.21. *Assume that the forms $Q(x,y) = ax^2 + bxy + cy^2$ and $P(x,y) = a'x^2 + b'xy + c'y^2$ have the same discriminant d and satisfy $gcd\left(a, a', \frac{b+b'}{2}\right) = 1$. Then there exists a unique integer B modulo $2aa'$ such that:*

$$\begin{aligned} B &\equiv b \ mod\ 2a, \\ B &\equiv b' \ mod\ 2a', \\ B^2 &\equiv d \ mod\ 4aa'. \end{aligned}$$

The proof of this result can be found in [8, p.43]. Now, we give Dirichlet's definition.

Definition 2.1.22. Let $Q(x,y) = ax^2 + bxy + cy^2$ and $P(x,y) = a'x^2 + b'xy + c'y^2$ be primitive positive definite forms of discriminant d satisfying $\gcd\left(a, a', \frac{b+b'}{2}\right) = 1$. We define the *Dirichlet composition* of $Q(x,y)$ and $P(x,y)$ to be

$$F(x,y) = aa'x^2 + Bxy + \frac{B^2 - d}{4aa'}y^2, \tag{2.1.19}$$

where B is given by Lemma 2.1.21.

Notice that it may happen that the condition $\gcd\left(a, a', \frac{b+b'}{2}\right) = 1$ of the above definition does not hold. In this case, we can replace one of the forms by a properly equivalent one which satisfies the condition.

Next, for a given discriminant d, we use $\mathcal{C}(d)$ to denote the set of classes of primitive positive definite forms of discriminant d. We have the following result which appears in [8, p.45].

Theorem 2.1.23. *Let d be a negative integer such that $d \equiv 0, 1 mod\ 4$. Then the Dirichlet composition (in Equation (2.1.19)) induces a well-defined binary operation on $\mathcal{C}(d)$ which makes $\mathcal{C}(d)$ into a finite abelian group of order the class number $h(d)$.*

Definition 2.1.24. We define the *form class group* to be the group $\mathcal{C}(d)$ as described in Theorem 2.1.23.

Obviously, one may think of what is the identity and the inverse element in the group $\mathcal{C}(d)$. The answer is that: the identity element is given by the class containing the principal form as defined in Equations (2.1.12) and (2.1.13); the inverse element of the class containing the form $ax^2 + bxy + cy^2$ is the class containing the form $ax^2 - bxy + cy^2$. In particular, the form $ax^2 - bxy + cy^2$ is called the *opposite* of $ax^2 + bxy + cy^2$.

2.2 Primitive Characters

We recall the following:

Definition 2.2.1. We say that a function χ from $\mathbb{Z}$ to $\mathbb{C}$ is a *Dirichlet character* if it has the following properties:

1. there exists a positive integer q such that $\chi(n) = \chi(n+q)$ for all n;
2. if $\gcd(n,q) > 1$ then $\chi(n) = 0$ and if $\gcd(n,q) = 1$ then $\chi(n) \neq 0$;
3. $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n.

In particular, a *principal character*, often denoted by χ_0, is defined by

$$\chi_0(n) = \begin{cases} 1 & \text{if } \gcd(n,q) = 1, \\ 0 & \text{if } \gcd(n,q) > 1. \end{cases} \tag{2.2.1}$$

Now, we are interested to a specific character called *primitive Dirichlet character.*

Definition 2.2.2. Let χ be a Dirichlet character modulo q and let ℓ be any positive divisor of q. We say that ℓ is an *induced modulus for* χ if we have

$$\chi(a) = 1 \text{ whenever } (a, q) = 1 \text{ and } a \equiv 1 \bmod \ell. \tag{2.2.2}$$

Definition 2.2.3. A Dirichlet character χ modulo q is called *primitive* modulo q if it has no induced modulus $\ell < q$. In other words, χ is *primitive* modulo q if and only if for any divisor $0 < \ell < q$, there is an integer $a \equiv 1 \bmod \ell$, $(a, q) = 1$ such that $\chi(a) \neq 1$.

As an example, we can observe that every nonprincipal character χ modulo a prime p is a primitive character modulo p. Also, if $q > 1$ the principal character is not primitive since it has 1 as an induced modulus.

Definition 2.2.4. Let χ be a Dirichlet character modulo q. We define the *conductor of* χ to be the smallest induced modulus ℓ for χ.

The importance of the conductor can be seen in the following:

Theorem 2.2.5. *[1, p.171] Every Dirichlet character χ modulo q can be expressed as*

$$\chi(n) = \psi(n)\chi_0(n), \qquad \text{for all } n,$$

where χ_0 is the principal character modulo q as in (2.2.1) *and ψ is a primitive character modulo the conductor of χ.*

We now state a result concerning the values of a real primitive character on a prime. The following can be found in [7, p.34].

Lemma 2.2.6. *Let χ be a real primitive character modulo m, and let p be an odd prime. Then we have*

$$\chi(p) = \left(\frac{\chi(-1)m}{p}\right), \tag{2.2.3}$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol.

In particular, we are interested to a special primitive Dirichlet character, namely the *Kronecker symbol* which is defined as follows.

Definition 2.2.7. Let n be nonzero integer with prime factorization $n = up_1^{e_1} \cdots p^{e_k}$. Let d be a fundamental discriminant (see Definition 2.1.4). We define the *Kronecker symbol* as

$$\left(\frac{d}{n}\right) = \left(\frac{d}{u}\right) \prod_{i=1}^{k} \left(\frac{d}{p_i}\right), \tag{2.2.4}$$

satisfying the following properties:

- $\left(\frac{d}{u}\right) = 1$ when $u = 1$,
- $\left(\frac{d}{-1}\right) = \begin{cases} 1 & \text{when } d > 0, \\ -1 & \text{when } d < 0, \end{cases}$
- $\left(\frac{d}{2}\right) = \begin{cases} 1 & \text{when } d \equiv 1 \bmod 8, \\ -1 & \text{when } d \equiv 5 \bmod 8, \end{cases}$
- $\left(\frac{d}{p_i}\right)$ is the Legendre symbol when p_i is odd,
- $\left(\frac{d}{p_i}\right) = 0$ when p_i divides d.

Proposition 2.2.8. *For any nonzero integer $d \equiv 0, 1 \bmod 4$, the Kronecker symbol is a Dirichlet character modulo $|d|$.*

As a particular character, for a given d, we often write

$$\chi_d(n) = \left(\frac{d}{n}\right) \tag{2.2.5}$$

for the primitive character obtained by the Kronecker symbol. We now prove the above proposition.
Proof of Proposition 2.2.8: For any $n_1, n_2 \in \mathbb{Z}^+$, we have $\chi_d(n_1 n_2) = \chi_d(n_1)\chi_d(n_2)$ which follows from the definition of Kronecker symbol in (2.2.4). Also, we have that $\chi_d(1) = (\frac{d}{1}) = 1$. Now, if $(n, |d|) = 1$, then we deduce by the property of the Kronecker symbol that $\chi_d(n) = \pm 1$ and $\chi_d(n) = 0$ when $(n, |d|) > 1$. □

Theorem 2.2.9. *If d is a fundamental discriminant, the Kronecker symbol $(\frac{d}{n})$ defines a real primitive character modulo $m = |d|$. Conversely, if χ is a real primitive character modulo m then $d = \chi(-1)m$ is a fundamental discriminant and $\chi(n) = (\frac{d}{n})$.*

Proof. We have already proved in Proposition 2.2.8 that χ_d is a character modulo $|d|$. Let us show that it is primitive. To do so, it is sufficient to consider a prime factor p of d and prove that the character cannot be defined modulo $\frac{d}{p}$. First, we suppose that $p \neq 2$. Let a be a quadratic nonresidue modulo p. As p is odd and d is a fundamental discriminant, we obtain that $\gcd\left(p, \frac{4|d|}{p}\right) = 1$. Using the Chinese Remainder Theorem, we can find $n > 0$ such that $n \equiv a \bmod p$ and $n \equiv 1 \bmod \frac{4|d|}{p}$; in particular, we have $n \equiv 1 \bmod 4$. Using Theorem 2.2.8 and the quadratic reciprocity law (for positive

odd), we may write

$$\begin{aligned}\left(\frac{d}{p}\right) &= \left(\frac{p}{n}\right)\left(\frac{d/p}{n}\right)\\ &= \left(\frac{p}{n}\right)\left(\frac{4d/p}{n}\right)\\ &= \left(\frac{p}{n}\right) = \left(\frac{n}{p}\right) = -1,\end{aligned}$$

which tells us that χ_d is not defined modulo d/p.

Now, we assume that $p = 2$, so that $d \equiv 8, 12 \bmod 16$, and we choose $n = 1 + |d|/2$. If $d \equiv 8 \bmod 16$, then $n \equiv 5 \bmod 8$ and $n \equiv 1 \bmod |d|/2$. We have that

$$\left(\frac{d}{n}\right) = \left(\frac{2}{n}\right)\left(\frac{d/2}{n}\right) = \left(\frac{2}{n}\right) = -1,$$

because $d/2 \equiv 0 \bmod 4$. On the other hand, if $d \equiv 12 \bmod 16$ then $n \equiv 7 \bmod 8$ and $n \equiv 1 \bmod d/4$, and we have

$$\left(\frac{d}{n}\right) = \left(\frac{-4}{n}\right)\left(\frac{-d/4}{n}\right) = \left(\frac{-4}{n}\right) = -1$$

because $-d/4 \equiv 1 \bmod 4$. Thus, $(\frac{d}{n})$ is not defined modulo $d/2$. Hence, we have shown that χ_d is a primitive character.

Conversely, let χ be a real primitive character modulo m and let p be an odd prime not dividing m. Recall from Lemma 2.2.3 that $\chi(p) = (\frac{d}{p})$ such that $d = \chi(-1)m$. As this is mutliplicative in p in both sides, then for any odd positive n we obtain $\chi(n) = (\frac{d}{n})$. Also, from the definition of the Kronecker symbol, we saw that $(\frac{d}{-1}) = sign(d) = \chi(-1)$. Thus we deduce that for any odd $n \in \mathbb{Z}$, $\chi(n) = (\frac{d}{n})$ holds. Next, let us prove that $d \equiv 0, 1 \bmod 4$. Since χ is of period $m = |d|$, by the properties of the Kronecker symbol, we have

$$1 = \chi(1 + 2d) = \left(\frac{d}{1 + 2d}\right).$$

Then, if $d \equiv 3 \bmod 4$ we will have

$$\begin{aligned}1 &= \left(\frac{-1}{1 + 2d}\right)\left(\frac{-d}{1 + 2d}\right)\\ &= (-1)^d\left(\frac{-d}{1}\right) = -1,\end{aligned}$$

which is a contradiction. Now if $d \equiv 2 \bmod 4$, we will have

$$\begin{aligned}1 &= \left(\frac{2}{1 + 2d}\right)\left(\frac{2d}{1 + 2d}\right)\\ &= \left(\frac{2}{1 + 2d}\right) = \left(\frac{2}{5}\right) = -1,\end{aligned}$$

which gives us also a contradiction. Let us show that $\chi(2) = (\frac{d}{2})$. Note that we may assume d to be odd, because otherwise we cannot have this equality. It follows that $\frac{d+1}{2}$ is odd, and so

$$\begin{aligned} 1 = \chi(d+1) &= \chi(2)\chi\Big((d+1)/2\Big) \\ &= \chi(2)\left(\frac{d}{(d+1)/2}\right) \\ &= \chi(2)\left(\frac{d}{d+1}\right)\left(\frac{d}{2}\right) \\ &= \chi(2)\left(\frac{d}{2}\right), \end{aligned}$$

which tells us that $\chi(2) = (\frac{d}{2})$. Note again that $\chi(n) = (\frac{d}{n})$ holds for any n by the multiplicativity property.

Now, as $d \equiv 0, 1 \bmod 4$, we may write $d = d_0 u^2$; with d_0 a fundamental discriminant. We remark that the two characters $(\frac{d}{n})$ and $(\frac{d_0}{n})$ take the same values for any n such that $\gcd(n, d) = 1$. Therefore, the character $(\frac{d}{n})$ is primitive if and only if $d = d_0$ meaning that d is a fundamental discriminant. □

Definition 2.2.10. Let d be a fundamental discriminant. We say that d is a *prime discriminant* if it is either equal to

$$-4,\ -8,\ 8 \text{ or } (-1)^{(p-1)/2}p \text{ for an odd prime } p. \tag{2.2.6}$$

Proposition 2.2.11. *Every fundamental discriminant d can be written uniquely as a product of prime discriminants.*

Proof. Recall that d is squarefree, so any power of an odd prime dividing d should be 1. It follows that

$$d = 2^{\alpha} \prod_{p \in \mathcal{P}} p,$$

where $\mathcal{P}$ denotes a finite set of odd primes. Hence, we may write

$$d = u 2^{\alpha} \prod_{p \in \mathcal{P}} (-1)^{(p-1)/2} p,$$

for some $u = \pm 1$. In addition, we note that the product is congruent to 1 modulo 4. Then, we have either $\alpha = 0$, $\alpha = 2$, or $\alpha = 3$. When $\alpha = 0$, we should have $u = 1$ because $d \equiv 1 \bmod 4$. When $\alpha = 2$, we should have $u = -1$ because otherwise $d/4$ becomes also a discriminant. And when $\alpha = 3$, we may have $u = \pm 1$, and so the factors ± 8. □

The main point is that the real primitive characters are identical to the Kronecker symbol $(\frac{d}{n})$, such that d is a product of prime discriminants. A more detailed explaination of this can be found in [8, p.40].

2.3 Quadratic Fields

In this section, we discuss a special case of number fields, namely the *quadratic fields*. These are number fields of degree 2 over $\mathbb{Q}$.

2.3.1 Invariants of quadratic fields

We begin with a general definition of an algebraic integer.

Definition 2.3.1. An element $\alpha \in K$ is called *algebraic integer* if it is the zero of certain monic polynomial with integer coefficients.

In the case of a quadratic field, we have the following: let $K = \mathbb{Q}(\alpha)$ where α is an algebraic interger, and is a zero of the monic polynomial $x^2 + ax + b$ with $a, b \in \mathbb{Z}$. And so, we have

$$\alpha = \frac{-a \pm \sqrt{a^2 - 4b}}{2}.$$

We set $a^2 - 4b = r^2 d$ where $r, d \in \mathbb{Z}$ and d is squarefree. Then, we may write

$$\alpha = \frac{-a \pm r\sqrt{d}}{2},$$

and now K is of the form $\mathbb{Q}(\sqrt{d})$. Indeed, we have shown the following:

Proposition 2.3.2. *Quadratic fields are precisely those of the form $\mathbb{Q}(\sqrt{d})$ where d is a squarefree integer.*

Definition 2.3.3. Let d be a squarefree integer. If $d > 0$ then $\mathbb{Q}(\sqrt{d})$ is called *real quadratic field*, and if $d < 0$ then $\mathbb{Q}(\sqrt{d})$ is called *imaginary quadratic field*.

Now, we are interested in the set of algebraic integers of the field $\mathbb{Q}(\sqrt{d})$, for d squarefree. We use O_K to denote this set where $K = \mathbb{Q}(\sqrt{d})$.

In the general case, more can be said about the set of algebraic integers O_K, where K is a number field. For example, it is known that O_K has a ring structure (see [15, p.96]). A more specific result is that the ring O_K is a free abelian group of rank $[K : \mathbb{Q}]$ the degree of K over $\mathbb{Q}$ (see [15, p.98]). As a standard reference on the study of number fields, a detailed discussion is provided in [19].

From now on, let K be a quadratic field that is $K = \mathbb{Q}(\sqrt{d})$ with d squarefree. In particular, the description of the ring of integers O_K is given as follows.

Theorem 2.3.4. *Let d be a squarefree integer. Then:*

$$O_K = \begin{cases} \mathbb{Z}[\sqrt{d}] & \text{if } d \not\equiv 1 \bmod 4, \\ \mathbb{Z}\left[\dfrac{1+\sqrt{d}}{2}\right] & \text{if } d \equiv 1 \bmod 4. \end{cases} \tag{2.3.1}$$

Proof. Any element $\alpha \in \mathbb{Q}(\sqrt{d})$ takes the form $\alpha = u + s\sqrt{d}$ with $u, v \in \mathbb{Q}$. So, we may write

$$\alpha = \frac{a + b\sqrt{d}}{c}, \tag{2.3.2}$$

such that $a, b, c \in \mathbb{Z}$ and no prime divides all a, b and c. Then, α is an algebraic integer if and only if the coeffficients of polynomial

$$\left(x - \frac{a + b\sqrt{d}}{c}\right)\left(x - \frac{a - b\sqrt{d}}{d}\right),$$

are integers, which is the minimum polynomial of α. Equivalenty, this says that

$$\frac{2a}{c} \in \mathbb{Z} \quad \text{and} \quad \frac{a^2 - b^2 d}{c^2} \in \mathbb{Z}\,. \tag{2.3.3}$$

Now, if we suppose that a prime p divides both a and c then the second condition in (2.3.3) says that b is divided by p as well. This cannot happen because of our hypothesis. Then, by using the first condition in (2.3.3), the possible values for c are 1 and 2. Asumme $c = 1$, then α is an integer of $\mathbb{Q}(\sqrt{d})$ in all the cases. If $c = 2$, then now a and b have to be both odd and the second condition in (2.3.3) holds. We have that

$$a^2 - b^2 d \equiv 0 \bmod 4.$$

Note that $4n^2 + 4n + 1 \equiv 1 \bmod 4$ is the square of an odd number $2n + 1$. That is, $a^2 \equiv 1 \bmod 4$ and $b^2 \equiv 1 \bmod 4$, and then $d \equiv 1 \bmod 4$. Conversely, if we have $d \equiv 1 \bmod 4$ then for a and b odd, α is an algebraic integer as the the two conditions in (2.3.3) hold. In fact, we have shown that when $d \equiv 1 \bmod 4$ we have $c = 2$ and a, b are odd; and if $d \not\equiv 1 \bmod 4$ then $c = 1$. □

One of the most basic invariants of a number field is the *norm*, *trace* and *discriminant*.

The norm and the trace of an element in the quadratic field is defined respectivelly as follows:

$$N(r + s\sqrt{d}) = r^2 - ds^2, \tag{2.3.4}$$

$$T(r + s\sqrt{d}) = 2r. \tag{2.3.5}$$

We want now to describe the notion of discriminant in the case of quadratic field. This can be done by using the two monomorphisms $\mathbb{Q}(\sqrt{d}) \longrightarrow \mathbb{C}$ given by

$$\sigma_1(r + s\sqrt{d}) = r + s\sqrt{d}, \tag{2.3.6}$$

$$\sigma_2(r + s\sqrt{d}) = r - s\sqrt{d}. \tag{2.3.7}$$

For any n-tuple of elements $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}(\sqrt{d})$, we define the discriminant of $\alpha_1, \ldots, \alpha_n$ to be

$$\operatorname{disc}(\alpha_1, \ldots, \alpha_n) = \Big(\det(\sigma_i(\alpha_j))\Big)^2, \tag{2.3.8}$$

that is, the square of the determinant of the matrix having $\sigma_i(\alpha_j)$ in the i^{th} row and j^{th} column $(i = 1, 2)$. Hence, we have:

Theorem 2.3.5. *Let d be a squarefree integer. Then the discriminant of a quadratic field $\mathbb{Q}(\sqrt{d})$ is*

$$\Delta = \begin{cases} 4d & \text{if } d \not\equiv 1 \bmod 4, \\ d & \text{if } d \equiv 1 \bmod 4. \end{cases} \tag{2.3.9}$$

Proof. First, we observe from Theorem 2.3.4 that an integral basis of $\mathbb{Q}(\sqrt{d})$ is given by $\{1, \tau\}$ where

$$\tau = \begin{cases} \sqrt{d} & \text{if } d \not\equiv 1 \bmod 4, \\ \dfrac{1+\sqrt{d}}{2} & \text{if } d \equiv 1 \bmod 4. \end{cases}$$

Using Equation (2.3.8), it follows that the discriminant of $\{1, \tau\}$ is

$$\begin{vmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{vmatrix}^2 = 4d,$$

when $d \not\equiv 1 \bmod 4$. Similarly, when $d \equiv 1 \bmod 4$ we get

$$\begin{vmatrix} 1 & \frac{1+\sqrt{d}}{2} \\ 1 & \frac{1-\sqrt{d}}{2} \end{vmatrix}^2 = d.$$

□

Remark 3. We have seen in Definition 2.1.4 what a fundamental discriminant is. According to the above Theorem 2.3.5, one may ask if there is a connection between the two notions. Indeed, a fundamental discriminant d is identical to $d = 1$ or $d = \Delta$ the discriminant of a quadratic field K. For this reason, we often say that Δ is a *fundamental quadratic discriminant*. Moreover, note that $\Delta \equiv 0, 1 \bmod 4$ and $K = \mathbb{Q}(\sqrt{\Delta})$. Hence, a quadratic field is determined by its discriminant Δ.

Next, we are going to see a generalization of the norm of an element in the field extension. More precisely, we consider ideal rather than element.

Definition 2.3.6. Let $\mathfrak{a}$ be an ideal of O_K. The *norm of the ideal* $\mathfrak{a}$, usually denoted by $N(\mathfrak{a})$, is the number of elements in the quotient $O_K / \mathfrak{a}$.

The following proposition provides some important properties of the norm of an ideal. The proof can be found in [17].

Proposition 2.3.7. *We have:*

1. *If $\mathfrak{a}, \mathfrak{b} \subset O_K$ are ideals, then $N(\mathfrak{a}\,\mathfrak{b}) = N(\mathfrak{a})N(\mathfrak{b})$.*
2. *Let $G = \{\sigma_1, \sigma_2\}$ (i.e. the Galois group of $\mathbb{Q}(\sqrt{d})\,/\,\mathbb{Q}$) where σ_i is defined as in (2.3.6) and (2.3.7). Then*

$$\prod_{\sigma \in G} \sigma(\mathfrak{a}) = (N(\mathfrak{a})), \tag{2.3.10}$$

   where $(N(\mathfrak{a}))$ is the principal ideal generated by $N(\mathfrak{a})$.
3. *Let $\alpha \in O_K$ and let $\mathfrak{a} = (\alpha)$ be the principal ideal generated by α. Let $N(\alpha)$ be the norm of α as in (2.3.4). Then $N(\mathfrak{a}) = |N(\alpha)|$.*

The above proposition shows that there is a relation between the norm of an element and the norm of an ideal. More specifically, the two notions of norms agree when we deal with principal ideals.

2.3.2 Arithmetic of $\mathbb{Q}(\sqrt{\Delta})$ and the ideal class group

Let K be a quadratic field. The set of ideals of O_K has a multiplicative structure. To make this precise, we need to introduce the notion of fractional ideals. Before that, let us state important properties of O_K.

Theorem 2.3.8. *[8, p.89] The ring of integers O_K is a Dedekind domain, which means that*

1. *O_K is integrally closed in K, that is, if $\alpha \in K$ satisfies a monic polynomial with coefficients in O_K then $\alpha \in O_K$.*
2. *O_K is Noetherian, that is, given a chain of ideals $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$ there exists an integer N such that $\mathfrak{a}_N = \mathfrak{a}_{N+1} = \cdots$.*
3. *Every nonzero prime ideal of O_K is maximal.*

The crucial point is given in the next result. In fact, a Dedekind domain has unique factorization (of elements) if and only if it is a principal ideal domain. More details about this latter fact can be found in [19, p.62].

Proposition 2.3.9. *[8, p.89] Every nonzero ideal $\mathfrak{a}$ in O_K can be written as*

$$\mathfrak{a} = \mathfrak{p}_1 \cdots \mathfrak{p}_r, \tag{2.3.11}$$

where $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ are prime ideals, and the decomposition is unique up to order. Furthermore, the $\mathfrak{p}_i$ are exactly the prime ideals of O_K containing $\mathfrak{a}$.

Next, we want to describe the arithmetic of O_K, that is *units* and *prime ideals* of O_K. For that, we consider the Kronecker symbol as in (2.2.5) such that d is now viewed as the discriminant of a quadratic field K. The following result provides us informations on the primes (i.e. prime ideals) of quadratic fields. The proof of the proposition can be found in [8, p.93].

Theorem 2.3.10. *Let K be a quadratic field of discriminant Δ, and let $\sigma_2 : \alpha \mapsto \alpha'$ be the nontrivial automorphism of K (see Equation (2.3.7)). Let p be prime in $\mathbb{Z}$.*

1. *If $\chi_\Delta(p) = 1$, then $pO_K = \mathfrak{p}\,\mathfrak{p}'$, where $\mathfrak{p} \neq \mathfrak{p}'$ are primes in O_K.*
2. *If $\chi_\Delta(p) = 0$, then $pO_K = \mathfrak{p}^2$ for some prime ideal $\mathfrak{p}$ of O_K.*
3. *If $\chi_\Delta(p) = -1$, then pO_K is prime in O_K.*

Furthermore, the primes above give all nonzero primes of O_K.

From this theorem, we can observe that if p is a rational prime, the norm of a prime ideal in O_K is p if $\chi_\Delta(p) = 0$ or $\chi_\Delta(p) = 1$, and p^2 if $\chi_\Delta(p) = -1$.

Now, we discuss the units of quadratic fields. For that, we use O_K^* to denote the unit group of the quadratic field K, that is the group of invertible elements of O_K.

Proposition 2.3.11. *[7, p.138] Let Δ be the discriminant of a quadratic field K. Then:*

1. *if $\Delta < -4$ we have $O_K^* = \{\pm 1\}$, that is $|O_K^*| = 2$;*
2. *if $\Delta = -4$ we have $O_K^* = \{\pm 1, \pm i\}$ (where $i^2 = -1$), that is $|O_K^*| = 4$;*
3. *if $\Delta = -3$ we have $O_K^* = \{\pm 1, \pm\rho, \pm\rho^2\} = \{(-\rho)^k \mid 0 \leq k \leq 5\}$ (where $\rho = \frac{1}{2}(-1 + \sqrt{-3})$ is a primitive cube root of unity), that is $|O_K^*| = 6$;*
4. *if $\Delta > 0$ there is $\varepsilon \in O_K$ (a fundamental unit) such that*

$$O_K^* = \{\pm 1\} \times \varepsilon^{\mathbb{Z}}, \tag{2.3.12}$$

that is, any $\mu \in O_K^*$ can be uniquely written as $\mu = \pm\varepsilon^k$ for some $k \in \mathbb{Z}$.

We observe that the first three cases are similar to the description in Equation (2.1.15).

We now introduce the notion of fractional ideals.

Definition 2.3.12. A *fractional ideal* $\mathfrak{a}$ of O_K is an O_K-submodule of K such that there exists some non-zero $c \in O_K$ with $c\,\mathfrak{a} \subset O_K$.

We know that O_K-submodules of O_K are ideals. In other words, the set $\mathfrak{b} = c\,\mathfrak{a}$ is an ideal of O_K, and we get $\mathfrak{a} = c^{-1}\,\mathfrak{b}$. Explicitly, the fractional ideals of O_K are subsets of K of the form $c^{-1}\,\mathfrak{b}$ where $\mathfrak{b}$ is an ideal of O_K and c is non-zero element of O_K.

We recall that, for any two ideals $\mathfrak{a}_1$ and $\mathfrak{a}_2$, the multiplication of ideals is given by

$$\mathfrak{a}_1\,\mathfrak{a}_2 = \left\{ \sum_{\text{finite}} xy \mid x \in \mathfrak{a}_1, y \in \mathfrak{a}_2 \right\}. \tag{2.3.13}$$

Next, we should note the following:

Remarks 1. 1. If we have $c\,\mathfrak{a} \subset O_K$ then $\bar{c}c\,\mathfrak{a} \subset O_K$. Note that the norm $N(c) = \bar{c}c = n \in \mathbb{Z}$, that is $n\,\mathfrak{a} \subset O_K$. Thus, in the second part of Definition 2.3.12, we may use the equivalent condition that *there exists non-zero integer n such that $n\,\mathfrak{a} \subset O_K$*. In this case, any fractional ideal is of the form $\mathfrak{a} = n^{-1}\,\mathfrak{b}$ for some $n \in \mathbb{Z}\setminus\{0\}$ with $\mathfrak{b}$ an ideal in O_K.

2. In general, an ideal is clearly a fractional ideal. Conversely, a fractional ideal $\mathfrak{a}$ is an ideal if and only if $\mathfrak{a} \subseteq O_K$. The product of fractional ideals gives a fractional ideal. This is because, if $\mathfrak{a}_1 = c_1^{-1}\,\mathfrak{b}_1$, $\mathfrak{a}_2 = c_2^{-1}\,\mathfrak{b}_2$ where $\mathfrak{b}_1, \mathfrak{b}_2$ are ideals and c_1, c_2 are non-zero elements of O_K, then the product is given by $\mathfrak{a}_1\,\mathfrak{a}_2 = (c_1c_2)^{-1}\,\mathfrak{b}_1\,\mathfrak{b}_2$.

The following result allows us to deduce that the set of fractional ideals has a group structure. It is taken from [8, p.90].

Proposition 2.3.13. *Let $\mathfrak{a}$ be a fractional ideal of O_K.*

1. *$\mathfrak{a}$ is invertible, that is there exists a fractional ideal $\mathfrak{b}$ of O_K such that $\mathfrak{a}\,\mathfrak{b} = O_K$.*

2. *$\mathfrak{a}$ can be written uniquely as a product $\prod_{i=1}^{s} \mathfrak{p}_i^{r_i}$ ($r_i \in \mathbb{Z}$) where $\mathfrak{p}_i$'s are distinct prime ideals of O_K.*

Now, we denote by $\mathcal{I}_K$ the set of all fractional ideals of O_K. Note that under the multiplication of ideals in (2.3.13) and the first part of Proposition 2.3.13, the multplication of fractional ideals is commutative, associative and every fractional ideal is invertible with O_K acting as an identity. Hence, it turns out that the set $\mathcal{I}_K$ is an abelian group under multiplication.

The most interesting subgroup of $\mathcal{I}_K$ is the subgroup of principal fractional ideals which is usually denoted by $\mathcal{P}_K$. These are ideals of the form $c^{-1}\,\mathfrak{b}$ where $\mathfrak{b}$ is a principal ideal of O_K. This means that principal fractional ideals are of the form αO_K for some $\alpha \in K$.

Definition 2.3.14. We define the *ideal class group* to be the quotient group

$$\mathcal{C}(O_K) = \mathcal{I}_K/\mathcal{P}_K. \tag{2.3.14}$$

The order of $\mathcal{C}(O_K)$ is called the *class number* and usually denoted by $h(O_K)$.

One of our goals in this chapter is to present the next theorem. It states a correspondence between fractional ideals and quadratic forms.

Theorem 2.3.15. *[8] Let K be a quadratic field of discriminant $\Delta < 0$. We have:*

1. *if $Q(x,y) = ax^2 + bxy + cy^2$ is a primitive positive definite quadratic form of discriminant Δ, then*

$$\left[a, \frac{-b+\sqrt{\Delta}}{2}\right] := \left\{ma + n\left(\frac{-b+\sqrt{\Delta}}{2}\right) \mid m, n \in \mathbb{Z}\right\} \tag{2.3.15}$$

   is an ideal of O_K.

2. *The map sending $Q(x,y)$ to $\left[a, \frac{-b+\sqrt{\Delta}}{2}\right]$ induces an isomorphism between the form class group $\mathcal{C}(\Delta)$ and the ideal class group $\mathcal{C}(O_K)$. Hence the order of $\mathcal{C}(O_K)$ is the class number $h(\Delta)$.*

Because of the isomorphism in the above theorem, we sometimes write the class number as $h(O_K)$ instead of $h(\Delta)$.

2.3.3 Ideals as $\mathbb{Z}$-modules

Here, we aim to determine the norm of an ideal in the ring of integers O_K where $K = \mathbb{Q}(\sqrt{\Delta})$ is a quadratic field. For doing so, we are going to classify all $\mathbb{Z}$-modules in the ring O_K. We start by recalling the notion of modules.

Definition 2.3.16. Let R be a commutative ring. An R-module M is an (additively) abelian group together with the map $R \times M \longrightarrow M$, $(r,m) \mapsto rm$ satisfying the following:

1. $1m = m$ for all $m \in M$,
2. $r(sm) = (rs)m$ for all $r, s \in R$ and $m \in M$,
3. $r(m+n) = rm + rn$ for all $r \in R$ and $m, n \in M$,
4. $(r+s)m = rm + sm$ for all $r, s \in R$ and $m \in M$.

Some classical examples are: abelian groups which are $\mathbb{Z}$-modules, subrings of a commutative ring R which are also $\mathbb{Z}$-modules. In particular, we are interested in the situation that an ideal of a commutative ring R is an R-module.

From Theorem 2.3.4, we know that $\{1, \tau\}$ is a basis of the ring of integers O_K where

$$\tau = \begin{cases} \sqrt{\Delta} & \text{if } \Delta \not\equiv 1 \bmod 4, \\ \frac{1+\sqrt{\Delta}}{2} & \text{if } \Delta \equiv 1 \bmod 4. \end{cases} \tag{2.3.16}$$

The following result gives us the classification of all $\mathbb{Z}$-modules in the ring O_K.

Proposition 2.3.17. *Let $M \subset O_K$ be a $\mathbb{Z}$-module in O_K. Then there exist nonnegative integers m, n and $a \in \mathbb{Z}$ such that*

$$M = [n, a + m\tau] := n\,\mathbb{Z} \oplus (a + m\tau)\,\mathbb{Z}, \tag{2.3.17}$$

where τ is as in Equation (2.3.16)

Proof. Let $G = \{s \in \mathbb{Z} \mid \text{ there is } r \in \mathbb{Z} \text{ and } r + s\tau \in M\}$ be a subgroup of $\mathbb{Z}$. For some $m \geq 0$, we know that G has the form $m\,\mathbb{Z}$. By construction, there exists $a \in \mathbb{Z}$ such that $a + m\tau \in M$. Since $M \cap \mathbb{Z}$ is a subgroup of $\mathbb{Z}$, thus there is some $n \geq 0$ such that $M \cap \mathbb{Z} = n\,\mathbb{Z}$.

We now want to show that $M = n\,\mathbb{Z} \oplus (a + m\tau)\,\mathbb{Z}$. The inclusion $\supseteq$ holds clearly. Let us show the other inclusion. Assume that $r + s\tau \in M$. As $s \in G$ we get $s = km$ for some $k \in \mathbb{Z}$. Then $r - ka = r + s\tau - k(a + m\tau) \in M \cap \mathbb{Z}$, that is $r - ka = ln$ for some $l \geq 0$. But then $r + s\tau = r - ka + k(a + m\tau) = ln + k(a + m\tau) \in n\,\mathbb{Z} \oplus (a + m\tau)\,\mathbb{Z}$. □

From Proposition 2.3.17, we can say that every element in M can be written uniquely as a $\mathbb{Z}$-linear combination of n and $a + m\tau$. Note that the elements n and $a + m\tau$ are called a *basis* of the $\mathbb{Z}$-module M.

In general, it may happen that some R-module does not have a basis. If an R-module does have a basis then it is called *free R-module* and the number of elements in a basis is called the *rank*. In other words, Proposition 2.3.17 says that all $\mathbb{Z}$-modules in O_K are free of rank less or equal to 2.

Moreover, we know that ideals in O_K can be viewed as O_K-modules. Since $\mathbb{Z} \subset O_K$ then every ideal in O_K is also a $\mathbb{Z}$-module. However, the converse is not true as for example $M = [1, 0] = \mathbb{Z}$ is a $\mathbb{Z}$-module in O_K but it is not an ideal (because we only have $(1) = O_K$).

Thus, one may ask:

QUESTION: Under what conditions is the $\mathbb{Z}$-module $M = [n, a + m\tau]$ an ideal? The answer is given by the following:

Proposition 2.3.18. *A non-zero $\mathbb{Z}$-module $M = [n, a + m\tau]$ is an ideal if and only if $m \mid n$, $m \mid a$ (hence $a = mb$ for some $b \in \mathbb{Z}$) and $n \mid m\,N(b + \tau)$.*

Proof. Since M is an ideal, $c \in M \cap \mathbb{Z}$ tells us that $c\tau \in M$, that is $c \in G$ where G is defined in the proof of Proposition 2.3.17. Hence, we have that $n\,\mathbb{Z} = M \cap \mathbb{Z} \subseteq G = m\,\mathbb{Z}$, thus m divides n.

Observe that $\tau^2 = x + y\tau$ for suitable $x, y \in \mathbb{Z}$. As M is an ideal, $a + m\tau \in M$ implies that $(a + m\tau)\tau = mx + (a + my)\tau \in M$, that is $a + my \in G$ by the definition of G. Then, $a + my$ is a multiple of m and hence m divides a.

Write $\alpha = a + m\tau = m(b + \tau)$. We have that $\alpha \in M$ implies $\alpha(b + \tau') \in M$. As $\frac{1}{m}N(\alpha) = m(b + \tau)(b + \tau') \in M \cap \mathbb{Z}$, we deduce $\frac{1}{m}N(b + \tau)$ is a multiple of n. □

Now, let R be a commutative ring. We have seen in Definition 2.3.6 that the norm of an ideal $\mathfrak{a}$ of R is given by the index of the additive group $\mathfrak{a}$ in the ring R. It is important to note that the same definition holds for $\mathbb{Z}$-submodules M of R. Indeed, the quotient R/M is an additive group and can be given a ring structure if M which happens to be an ideal. We have:

Definition 2.3.19. Let M be a $\mathbb{Z}$-submodule of a ring R. The norm of M is defined as

$$N(M) = |R/M| \tag{2.3.18}$$

the cardinality of the quotient group R/M.

Note that, in general, it may happen that $N(M)$ is not finite. For example, take $M = [1, 0] = \mathbb{Z}$, $R = O_K$ and hence $R/M = \{b \mid b \in \mathbb{Z}\}$. In particular, $|R/M| = \infty$.

On the other hand, this cannot happen when M is a $\mathbb{Z}$-module of rank 2 (i.e. when M is an ideal). Note that a $\mathbb{Z}$-module $M = [n, a + m\tau]$ in O_K has rank 2 if and only if $mn \neq 0$. In particular, the first part of Theorem 2.3.15 tells us that the $\mathbb{Z}$-module $[a, \frac{-b+\sqrt{\Delta}}{2}]$ has rank 2. In such a case, we say that $[a, \frac{-b+\sqrt{\Delta}}{2}]$ is a *full module*.

Next, we have the following proposition which gives us the norm of an ideal in O_K.

Proposition 2.3.20. *Let $M = [n, a + m\tau]$ be a full $\mathbb{Z}$-module in O_K. Then*

$$S = \{r + s\tau : 0 \leq r < n \text{ and } 0 \leq s < m\}$$

is a complete residue system modulo M in O_K, and in particular $N(M) = mn$.

Proof. First, we show that every $x + y\tau \in O_K$ is congruent modulo M to an element of S. Write $y = mq + s$ for some $q \in \mathbb{Z}$ and $0 \leq s < m$. Then $x + y\tau - q(a + m\tau) = x' + s\tau$ for some integer x', hence $x + y\tau \equiv x' + s\tau \bmod M$. Now, we write $x' = nq' + r$ for $q' \in \mathbb{Z}$ and $0 \leq r < n$. Then $x' + s\tau \equiv r + s\tau \bmod M$.

Our claim is that the elements of S are pairwise incongruent modulo M. Suppose that $r + s\tau \equiv r' + s'\tau \bmod M$ for $0 \leq r, r' < n$ and $0 \leq s, s' < m$. Then $r - r' + (s - s')\tau \in M$ implies that $s - s' \in m\,\mathbb{Z}$ and $r - r' \in n\,\mathbb{Z}$. Hence $r = r'$ and $s = s'$. □

Hence, when K is an imaginary quadratic field of (fundamental) discriminant Δ, we get:

Proposition 2.3.21. *Every ideal $\mathfrak{a}$ of O_K can be uniquely represented as*

$$\mathfrak{a} = u\left[a, \frac{-b + \sqrt{\Delta}}{2}\right], \tag{2.3.19}$$

for some positive integers u and a such that the integer b is determined by

$$-a < b \leq a \quad \text{and} \quad b^2 \equiv \Delta \bmod 4a. \tag{2.3.20}$$

Moreover, the norm of $\mathfrak{a}$ is given by $N(\mathfrak{a}) = u^2 a$.

Proof. Let $\mathfrak{a}$ be an ideal of O_K. By Proposition 2.3.18, $\mathfrak{a}$ is of the form

$$\mathfrak{a} = \left[ua, u\left(\frac{-b+\sqrt{\Delta}}{2}\right)\right] \quad \text{with} \quad a \mid N\left(\frac{-b+\sqrt{\Delta}}{2}\right), \tag{2.3.21}$$

for some positive integers u and a. Hence, $\mathfrak{a} = u[a, \frac{-b+\sqrt{\Delta}}{2}]$. From Proposition 2.3.20, we get $N(\mathfrak{a}) = u^2 a$. By the second part of Equation (2.3.21), the condition

$$b^2 \equiv \Delta \bmod 4a \tag{2.3.22}$$

holds since we have that

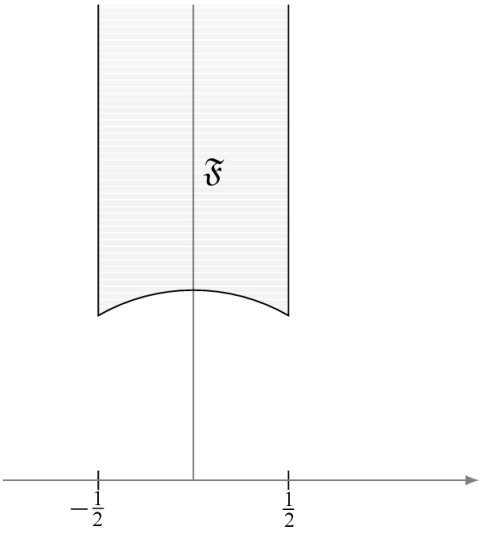
$$N\left(\frac{-b+\sqrt{\Delta}}{2}\right) = \frac{b^2-\Delta}{4}, \tag{2.3.23}$$

and because a divides this norm then Equation(2.3.22) is obtained.

Write $\sigma_i = \frac{-b_i+\sqrt{\Delta}}{2}$. If an ideal $\mathfrak{a}$ is written into two forms then we have $\mathfrak{a} = u_1[a_1, \sigma_1] = u_2[a_2, \sigma_2]$ with positive integers u_i, a_i for $i = 1, 2$. So, we get

$$u_1 a_1 \left[1, \frac{\sigma_1}{a_1}\right] = u_2 a_2 \left[1, \frac{\sigma_2}{a_2}\right]. \tag{2.3.24}$$

Denote $\omega_i = \frac{\sigma_i}{a_i}$ for $i = 1, 2$. The previous equation means that $u_1 a_1[1, \omega_1] = u_2 a_2[1, \omega_2]$. Now, we consider the standard fundamental domain $\mathfrak{F}$ defined as in the following figure.



More precisely,

$$\mathfrak{F} = \Big\{ w \in \mathbb{C} \mid \mathrm{Im}(w) > 0,\ -\frac{1}{2} \leq \mathrm{Re}(w) < \frac{1}{2} \quad \text{and} \quad |w| \geq 1 \Big\}. \tag{2.3.25}$$

For $i = 1, 2$, observe that

$$-\frac{1}{2} \leq \mathrm{Re}(\omega_i) < \frac{1}{2},\ \ \mathrm{Im}(\omega_i) > 0 \text{ and } |\omega_i| \geq 1, \tag{2.3.26}$$

because $-a_i < b_i \leq a_i$. In other words, we have that $\omega_1, \omega_2 \in \mathfrak{F}$. In terms of lattice, Equation (2.3.24) says that there exists an unimodular transformation $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\frac{p\omega_1 + q}{r\omega_1 + s} = \omega_2. \tag{2.3.27}$$

Since ω_i lies in the fundamental domain for the modular group $\mathrm{SL}_2(\mathbb{Z})$, then $\omega_1 = \omega_2$. In particular, we have $a_1 = a_2$ and $b_1 = b_2$. By Equation (2.3.24), we also get $u_1 = u_2$. Hence, $\mathfrak{a}$ is uniquely of the form $u[a, \frac{-b+\sqrt{\Delta}}{2}]$ for some nonzero integers u and a such that $-a < b \leq a$ and Equation (2.3.22) holds. □

More detailed discussion about the representation of ideals in quadratic extensions can be found in the textbook [6, Chap.2, Sec.2.6].

2.3.4 Dedekind zeta function for quadratic fields

Following our discussion on the ideal of the ring of integers O_K, we introduce a special function which is known as the generalization of the Riemann zeta function $\zeta(s)$ $(\mathrm{Re}(s) > 1)$, namely the *Dedekind zeta function* $\zeta_K(s)$ for a number field K.

Definition 2.3.22. Let K be a number field. We define the *Dedekind zeta function* for K, usually denoted by $\zeta_K(s)$, as

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s} = \prod_{\mathfrak{p}} (1 - N(\mathfrak{p})^{-s})^{-1}, \tag{2.3.28}$$

for $\mathrm{Re}(s) > 1$ and the sum runs over all nonzero ideals $\mathfrak{a}$ in O_K. Note that the second equality is the Euler product for $\zeta_K(s)$ such that the product is over all prime ideals $\mathfrak{p}$ of O_K.

In particular, we obtain the Dedekind zeta function associated to the quadratic field $\mathbb{Q}(\sqrt{\Delta})$. There are many results that can be said about $\zeta_{\mathbb{Q}(\sqrt{\Delta})}(s)$, but we are only interested on the theorem below which will be used later.

Proposition 2.3.23. *For* $\mathrm{Re}(s) > 1$ *we have*

$$\zeta_K(s) = \zeta(s) L(s, \chi_\Delta). \tag{2.3.29}$$

Proof. The proof is mainly based on the decomposition of a prime $p \in \mathbb{Z}$ into prime ideals of O_K. From our earlier discussion, we know what the norm of a prime ideal $\mathfrak{p} \in O_K$ is (see Theorem 2.3.10). Also, it should be noted that in the case $(\frac{\Delta}{p}) = 1$, we obtain two prime ideals since the rational prime p spilts. Now, for $\mathrm{Re}(s) > 1$, we may write

$$\begin{aligned}
\zeta_K(s) &= \prod_{\mathfrak{p}} (1 - N(\mathfrak{p})^{-s})^{-1}, \\
&= \prod_{p,(\frac{\Delta}{p})=1} (1-p^{-s})^{-2} \times \prod_{p,(\frac{\Delta}{p})=-1} (1-p^{-2s})^{-1} \times \prod_{p,(\frac{\Delta}{p})=0} (1-p^{-s})^{-1}, \\
&= \prod_{p} (1-p^{-s})^{-1} \left[\prod_{p,(\frac{\Delta}{p})=1} (1-p^{-s})^{-1} \times \prod_{p,(\frac{\Delta}{p})=-1} (1+p^{-s})^{-1} \right], \\
&= \prod_{p} (1-p^{-s})^{-1} \prod_{p} \left[1 - \left(\frac{\Delta}{p} \right)^{-s} p^{-s} \right]^{-1}.
\end{aligned}$$

By the Euler product of the Riemann zeta function and the L-function, we finally get that

$$\zeta_K(s) = \zeta(s) L(s, \chi_\Delta),$$

where χ_Δ is the Kronecker symbol as in (2.2.5). □

2.4 Analytic Class Number Formula

In this section, we aim to show the Dirichlet's class number formula in the negative case, that is, when Δ is a negative fundamental discriminant. For doing so, we will consider three cases depending on the value of Δ. It should be noted that this formula can be viewed as a way of seeing the nonvanishing of the L-function at $s = 1$.

We begin by stating the general result.

Theorem 2.4.1. *Let $\Delta < 0$ be fundamental quadratic discriminant. We have*

$$L(1, \chi_\Delta) = \frac{2\pi h(\Delta)}{\omega \sqrt{|\Delta|}}, \tag{2.4.1}$$

where $h(\Delta)$ denotes the class number, χ_Δ is defined as in Equation (2.2.5), and ω is given in (2.1.15).

As we observed earlier in Section 2.1, the two values where $\Delta = -4$ and $\Delta = -3$ have both a particular number of automorphs which is given in Proposition 2.1.18. Equivalently, this can also be seen from Proposition 2.3.11. So, we would like to treat them independently.

2.4.1 Case $\Delta < -4$

We concentrate in the case where $\Delta < -4$, so the general statement is as follows when $\Delta < -4$.

Theorem 2.4.2. *For $\Delta < -4$ a fundamental discriminant, we have*

$$L(1, \chi_\Delta) = \frac{\pi h(\Delta)}{\sqrt{|\Delta|}}, \tag{2.4.2}$$

where $h(\Delta)$ denotes the class number and χ_Δ is defined in (2.2.5).

As one can see, this theorem relates the L-function with the class number. To prove it, we have to introduce a function depending on the equivalance class of quadratic forms, namely the *Epstein zeta function*.

Definition 2.4.3. Let $F(x, y)$ be a quadratic form of discriminant Δ. We define the *Epstein zeta function* as a function defined for $\mathrm{Re}(s) > 1$ by

$$\zeta(s, F) = \frac{1}{2} \sum_{(x,y)} F(x, y)^{-s}, \tag{2.4.3}$$

where the sum runs over all pairs of integers $(x, y) \neq (0, 0)$.

Note that since equivalent forms represent the same values, we get the same sum if we replace the form F by an equivalent form. In other words, the function $\zeta(s, F)$ only depends on the class of the form $F(x, y)$.

Let us begin with some discussion on the index of the sum (2.4.3). We remark that we were summing over all nonzero pairs of integers (x, y). So it makes sense to factor out the greatest common divisor of the each individual pair (x, y), say m. Now, the sum can be changed to a sum over all possible greatest common divisor m and a sum over all pairs (u, v) with u and v relatively prime. More precisely, we have that

$$\begin{aligned}
\sum_{(x,y)} F(x, y)^{-s} &= \sum_{m=1}^{\infty} \sum_{(u,v)} F(mu, mv)^{-s}, \\
&= \sum_{m=1}^{\infty} \sum_{(u,v)} \left(m^2 F(u, v)\right)^{-s}, \\
&= \sum_{m=1}^{\infty} m^{-2s} \sum_{(u,v)} F(u, v)^{-s} \\
&= \zeta(2s) \sum_{(u,v)} F(u, v)^{-s}.
\end{aligned}$$

We are now focussing on the forms $F(u, v)$ where u and v are relatively prime. Remember that this form may represent some integer, so we gather all

forms $F(u,v)$ which take the same value, say n. From Definition 2.1.19, we obtain $r_F(n)$ of them. In fact, our definition of $\zeta(s,F)$ in (2.4.3) becomes

$$\zeta(s,F) = \zeta(2s)\sum_{n=1}^{\infty} r_{[F]}(n)n^{-s}. \tag{2.4.4}$$

Summing over all $h(\Delta)$ equivalence classes of forms, we obtain

$$\begin{aligned}\zeta(s,\Delta) &= \sum_{[F]\in\mathcal{C}(\Delta)} \zeta(s,F), &(2.4.5)\\ &= \zeta(2s)\sum_{n=1}^{\infty}\left(\sum_{[F]} r_{[F]}(n)\right)n^{-s} &(2.4.6)\\ &= \zeta(2s)\sum_{n=1}^{\infty} r_{\Delta}(n)n^{-s}. &(2.4.7)\end{aligned}$$

Lemma 2.4.4. *We have an Euler product*

$$\sum_{n=1}^{\infty} r_{\Delta}(n)n^{-s} = \prod_{p}\frac{1+p^{-s}}{1-\chi_{\Delta}(p)p^{-s}}. \tag{2.4.8}$$

Proof. By considering the right hand side, we have that

$$\begin{aligned}\frac{1+p^{-s}}{1-\chi_{\Delta}(p)p^{-s}} &= (1+p^{-s})\sum_{k=0}^{\infty}\chi_{\Delta}(p)^k p^{-ks},\\ &= \sum_{k=0}^{\infty}\chi_{\Delta}(p)^k p^{-ks} + \sum_{k=0}^{\infty}\chi_{\Delta}(p)^k p^{-(k+1)s},\\ &= 1+p^{-s}+\chi_{\Delta}(p)p^{-s}+\chi_{\Delta}(p)p^{-2s}+\chi_{\Delta}(p)^2p^{-2s}+\chi_{\Delta}(p)^2p^{-3s}+\cdots,\\ &= 1+\Big[1+\chi_{\Delta}(p)\Big]p^{-s}+\sum_{k=2}^{\infty}\Big[\chi_{\Delta}(p)^{k-1}+\chi_{\Delta}(p)^k\Big]p^{-ks}.\end{aligned}$$

One can see that the above calculation depends on the value of $\chi_{\Delta}(p)$. In fact, we have

$$\frac{1+p^{-s}}{1-\chi_{\Delta}(p)p^{-s}} = \begin{cases}1+2\sum_k p^{-ks} & \text{if } \chi_{\Delta}(p)=1,\\ 1+p^{-s} & \text{if } \chi_{\Delta}(p)=0,\\ 1 & \text{if } \chi_{\Delta}(p)=-1,\end{cases} \tag{2.4.9}$$

Now, we multiply all the factors and notice that the coefficient of n^{-s} is $r_{\Delta}(n)$ which is from Theorem 2.1.20. □

Lemma 2.4.5. *We have that*

$$\prod_{p}\frac{1+p^{-s}}{1-\chi_{\Delta}(p)p^{-s}} = \zeta(2s)^{-1}L(s,\chi_{\Delta})\zeta(s). \tag{2.4.10}$$

Proof. Considering the left hand side, we have for $\mathrm{Re}(s) > 1$ that

$$\begin{aligned}\prod_p \frac{1+p^{-s}}{1-\chi_\Delta(p)p^{-s}} &= \prod_p \frac{(1+p^{-s})(1-p^{-s})}{(1-\chi_\Delta(p)p^{-s})(1-p^{-s})},\\ &= \prod_p \frac{1-p^{-2s}}{(1-\chi_\Delta(p)p^{-s})(1-p^{-s})},\\ &= \prod_p \left(1-p^{-2s}\right)\prod_p \left(1-\chi_\Delta(p)p^{-s}\right)^{-1}\prod_p \left(1-p^{-s}\right)^{-1},\\ &= \zeta(2s)^{-1}L(s,\chi_\Delta)\zeta(s).\end{aligned}$$

□

So now, we get a new way of writting the function $\zeta(s,\Delta)$ which is given as follows:

Theorem 2.4.6. *We have*

$$\zeta(s,\Delta) = \zeta(s)L(s,\chi_\Delta). \tag{2.4.11}$$

Proof. Using Lemma 2.4.4 and 2.4.5, it follows that

$$\begin{aligned}\zeta(s,\Delta) &= \zeta(2s)\sum_{n=1}^{\infty} r_\Delta(n)n^{-s},\\ &= \zeta(2s)\zeta(2s)^{-1}L(s,\chi_\Delta)\zeta(s),\\ &= \zeta(s)L(s,\chi_\Delta).\end{aligned}$$

□

Note that this theorem is similar with Theorem 2.3.23. We are now focussing at the behavior of $\zeta(s,\Delta)$ at $s=1$. To do so, we need the following function;

Definition 2.4.7. We define

$$\Theta(t,F) = \sum_{(x,y)} \exp\left[\frac{-2\pi t F(x,y)}{\sqrt{|\Delta|}}\right], \tag{2.4.12}$$

with $\Delta < 0$ the discriminant of the form $F(x,y)$.

Remark 4. The function $\Theta(t,F)$ depends on the class of the forms $F(x,y)$. Moreover, from [24, p.314], $\Theta(t,F)$ is symmetric in the sense that

$$\Theta(t^{-1},F) = t\Theta(t,F'),$$

for some form $F'(x,y) = ax^2 - bxy + cy^2$ and $F(x,y) = ax^2 + bxy + cy^2$.

Another definition that we need is also the following:

Definition 2.4.8. Let $F(x,y)$ be a quadratic form of discriminant Δ. We define

$$\Lambda(s,F) = 2|\Delta|^{\frac{s}{2}}(2\pi)^{-s}\Gamma(s)\zeta(s,F), \tag{2.4.13}$$

and also

$$\Lambda(s,\Delta) = \sum_{[F]} \Lambda(s,F). \tag{2.4.14}$$

It follows from Theorem 2.4.6 that

$$\Lambda(s,\Delta) = 2|\Delta|^{\frac{s}{2}}(2\pi)^{-s}\Gamma(s)\zeta(s)L(s,\chi_\Delta). \tag{2.4.15}$$

Proposition 2.4.9. *The following property of $\Lambda(s,F)$ holds:*

1.
$$\Lambda(s,F) = \int_0^\infty (\Theta(t,F)-1)t^s\frac{dt}{t}, \tag{2.4.16}$$

2. *(Symmetricity)*
$$\Lambda(s,F) = \frac{1}{s-1} - \frac{1}{s} + \int_1^\infty (\Theta(t,F)-1)t^s\frac{dt}{t} + \int_1^\infty (\Theta(t,F')-1)t^{1-s}\frac{dt}{t}. \tag{2.4.17}$$

Proof. • Recall that our original definition of $\zeta(s,F)$ is given by

$$\frac{1}{2}\sum_{(x,y)} F(x,y)^{-s}, \qquad \text{for } \mathrm{Re}(s) > 1.$$

Now for an integer n, we consider the integral

$$\int_0^\infty \exp[-\pi n^2 t]t^s\frac{dt}{t} = \frac{1}{n^{2s}}\Gamma(s)\pi^{-s}. \tag{2.4.18}$$

By summing both side of (2.4.18) by all pairs (x,y) and replace n^2 by $\frac{2F(x,y)}{\sqrt{|\Delta|}}$, we get that

$$\sum_{(x,y)}\int_0^\infty \exp\left[\frac{-2\pi t F(x,y)}{\sqrt{|\Delta|}}\right]t^s\frac{dt}{t} = \sum_{(x,y)}\frac{|\Delta|^{\frac{s}{2}}}{2^s F(x,y)^s}\Gamma(s)\pi^{-s} = \sum_{(x,y)}|\Delta|^{\frac{s}{2}}F(x,y)^{-s}\Gamma(s)(2\pi)^{-s}. \tag{2.4.19}$$

So, we have

$$\begin{aligned}\Lambda(s,F) &= 2|\Delta|^{\frac{s}{2}}(2\pi)^{-s}\Gamma(s)\zeta(s,F),\\ &= 2|\Delta|^{\frac{s}{2}}(2\pi)^{-s}\Gamma(s)\frac{1}{2}\sum_{(x,y)}F(x,y)^{-s},\\ &= |\Delta|^{\frac{s}{2}}(2\pi)^{-s}\Gamma(s)\sum_{(x,y)}F(x,y)^{-s}.\end{aligned}$$

And as one can observe, this is exactly as in (2.4.19). Then, it follows that

$$\Lambda(s,F)=\sum_{(x,y)}\int_0^\infty \exp\left[\frac{-2\pi tF(x,y)}{\sqrt{|\Delta|}}\right]t^s\frac{dt}{t}. \tag{2.4.20}$$

Comparing this result with the definition of the function $\Theta(t,F)$, we may write

$$\Lambda(s,F)=\int_0^\infty(\Theta(t,F)-1)t^s\frac{dt}{t}, \tag{2.4.21}$$

since we only consider a nonzero pair (x,y).

• The proof of the symmetricity property is based on the property of the function $\Theta(t,F)$ given in Remark 4. Using the result from the first point and by splitting the integral, we find that

$$\Lambda(s,F)=\int_0^1(\Theta(t,F)-1)t^s\frac{dt}{t}+\int_1^\infty(\Theta(t,F)-1)t^s\frac{dt}{t}.$$

Let us focus on the first integral which is defined between 0 and 1. We set $f(t,F)=\Theta(t,F)-1$ so that $\Theta(t,F)=1+f(t,F)$. Recall that there is a property of $\Theta(t,F)$ given by $\Theta(t,F')=t^{-1}\Theta(t^{-1},F)$ as in Remark 4. Then,

$$\begin{aligned}\Theta(t,F')&=1+f(t,F'),\\&=t^{-1}\Big(1+f(t^{-1},F)\Big),\end{aligned}$$

which allows us to write $f(t,F')=t^{-1}\Big(1+f(t^{-1},F)\Big)-1$. Thus we obtain the following

$$\begin{aligned}\int_0^1 f(t,F)t^s\frac{dt}{t}&=\int_0^1\left[t^{-1}\left(1+f(t^{-1},F')\right)-1\right]t^s\frac{dt}{t},\\&=\int_0^1(t^{s-1}-t^s)\frac{dt}{t}+\int_0^1 t^{-1}f(t^{-1},F')t^s\frac{dt}{t},\\&=\int_0^1(t^{s-2}-t^{s-1})\,dt+\int_0^1 t^{-1}f(t^{-1},F')t^s\frac{dt}{t},\\&=\frac{t^{s-1}}{s-1}-\frac{t^s}{s}\Big|_0^1+\int_0^1 t^{-1}f(t^{-1},F')t^s\frac{dt}{t},\\&=\frac{1}{s-1}-\frac{1}{s}+\int_0^1 t^{-1}f(t^{-1},F')t^s\frac{dt}{t}.\end{aligned}$$

The remaining integral can be treated as follows: we set $z=t^{-1}$ so that $dz=-t^{-2}dt$, that is $\frac{dz}{z}=-\frac{dt}{t}$. Then,

$$\int_0^1 t^{-1}f(t^{-1},F')t^s\frac{dt}{t}=\int_1^\infty zf(z,F')z^{-s}\frac{dz}{z}=\int_1^\infty f(z,F')z^{1-s}\frac{dz}{z}, \tag{2.4.22}$$

which can be written again as

$$\int_1^\infty f(t,F')t^{1-s}\frac{dt}{t}.$$

Thus, we finally have

$$\Lambda(s,F)=\frac{1}{s-1}-\frac{1}{s}+\int_1^\infty f(t,F')t^{1-s}\frac{dt}{t}+\int_1^\infty(\Theta(t,F)-1)t^s\frac{dt}{t}. \quad (2.4.23)$$

□

We deduce from the second part of the above proposition that

$$\Lambda(s,F)=\frac{1}{s-1}+O(1), \quad \text{as } s\to 1. \quad (2.4.24)$$

Now, using Definition 2.4.14, we find that

$$\Lambda(s,\Delta)=\frac{h(\Delta)}{s-1}+O(1), \quad \text{as } s\to 1, \quad (2.4.25)$$

such that here the class number $h(d)$ appears. We know that at $s=1$, we have

$$\zeta(s)=\frac{1}{s-1}+O(1), \quad \text{as } s\to 1. \quad (2.4.26)$$

Using Equations (2.4.15) and (2.4.26), we get

$$\Lambda(s,\Delta)=\frac{2\sqrt{|\Delta|}(2\pi)^{-1}\Gamma(1)L(1,\chi_\Delta)}{s-1}+O(1), \quad \text{as } s\to 1, \quad (2.4.27)$$

In fact, we take Equation (2.4.26) and multiply this with the remaining right hand side of Equation (2.4.15) at $s=1$. Note that all the functions in Equation (2.4.15) are well behaved at $s=1$.

Finally, to deduce the class number formula, we identify the two equations in (2.4.25) and in (2.4.27), which gives us

$$L(1,\chi_\Delta)=\frac{\pi h(d)}{\sqrt{|\Delta|}}.$$

2.4.2 Case $\Delta=-4$

Next, we are going to evaluate the series $L(s,\chi_{-4})$ at $s=1$. Our approach consists of using a specific partial sums and by identification, we deduce the value of $L(1,\chi_{-4})$.

Note that $d = -4$ is a fundamental discriminant. Recall that the character χ_{-4} is given as follows: $\chi_{-4}(n) = 0$ when n is even, and for an odd n we have

$$\begin{aligned}\chi_{-4}(n) &= \left(\frac{-4}{n}\right) \\ &= \left(\frac{-1}{n}\right)\left(\frac{4}{n}\right) \\ &= (-1)^{(n-1)/2}.\end{aligned}$$

Then, we get that

$$\chi_{-4}(n) = \begin{cases} 1 & \text{if } n \equiv 1 \bmod 4, \\ -1 & \text{if } n \equiv 3 \bmod 4. \end{cases} \tag{2.4.28}$$

It follows that the Dirichlet L-functions associated to χ_{-4} is given by

$$L(s, \chi_{-4}) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \frac{1}{9^s} - \frac{1}{11^s} + \frac{1}{13^s} - \cdots . \tag{2.4.29}$$

We would like to evaluate this sums. Clearly, we known that

$$\sum_{n=0}^{2N} (-1)^n t^{2n} = \frac{1}{1+t^2} + \frac{t^{4N+2}}{1+t^2}. \tag{2.4.30}$$

We now integrate both side of (2.4.30) between 0 and 1. Thus,

$$\int_0^1 \left[\sum_{n=0}^{2N} (-1)^n t^{2n}\right] dt = \sum_{n=0}^{2N} \int_0^1 (-1)^n t^{2n} dt = \sum_{n=0}^{2N} (-1)^n \frac{t^{2n+1}}{2n+1}\bigg|_0^1 = \sum_{n=0}^{2N} \frac{(-1)^n}{2n+1}.$$

On the other hand, the right hand side gives us

$$\int_0^1 \frac{dt}{1+t^2} + \int_0^1 \frac{t^{4N+2}}{1+t^2}\, dt.$$

Note that the first expression is equal to

$$\int_0^1 \frac{dt}{1+t^2} = \arctan(1) = \frac{\pi}{4}, \tag{2.4.31}$$

and the second expression is bounded by

$$\int_0^1 \frac{t^{4N+2}}{1+t^2}\, dt < \int_0^1 t^{4N+2}\, dt = \frac{t^{4N+3}}{4N+3}\bigg|_0^1 = \frac{1}{4N+3},$$

because we have $0 < t < 1$. Hence, by tending N to ∞, we obtain that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}. \tag{2.4.32}$$

Plugging $s = 1$ in the expression of $L(s, \chi_{-4})$ in (2.4.29), we get that

$$L(1, \chi_{-4}) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \frac{1}{13} - \cdots = \sum_{n=0}^{2N} \frac{(-1)^n}{2n+1}. \quad (2.4.33)$$

Therefore, we deduce that

$$L(1, \chi_{-4}) = \frac{\pi}{4}. \quad (2.4.34)$$

As we can see, Equation (2.4.34) is a special case of Theorem 2.4.1 in the sense that, for $\Delta = -4$,

$$L(1, \chi_{-4}) = \frac{\pi}{4} = \frac{2\pi h(-4)}{\omega\sqrt{4}}. \quad (2.4.35)$$

It follows that $h(-4) = 1$ since $\omega = 4$. The fact that $h(\Delta) = 1$ for $\Delta = -4$ is a direct consequence of Theorem 2.1.10. Indeed, Condition (2.1.10) allows us to enumerate all the reduced forms, thus, to compute the class number. Observe that for a reduced form $ax^2 + bxy + cy^2$ of discriminant Δ, we have

$$|b| \leq \sqrt{\frac{|\Delta|}{3}} \quad \text{and} \quad b \equiv \Delta \bmod 2. \quad (2.4.36)$$

Moreover, we also have

$$4ac = b^2 - \Delta. \quad (2.4.37)$$

When $\Delta = -4$, Equation (2.4.36) gives us $b = 0$. By Equation (2.4.37), the only possible choice of a should divide $\frac{1}{4}(b^2 - \Delta) = 1$. So, we have $a = 1$ and then $c = 1$. Hence, $x^2 + y^2$ is the only reduced form. Therefore, $h(-4) = 1$.

2.4.3 Case $\Delta = -3$

We recall that the Dirichlet character χ_{-3} is defined as follows

$$\chi_{-3}(n) = \begin{cases} 0 & \text{if } n \equiv 0 \bmod 3, \\ 1 & \text{if } n \equiv 1 \bmod 3, \\ -1 & \text{if } n \equiv 2 \bmod 3. \end{cases}$$

As we did before, let us write the definition of $L(s, \chi_{-3})$, such as

$$L(s, \chi_{-3}) = 1 - \frac{1}{2^s} + \frac{1}{4^s} - \frac{1}{5^s} + \frac{1}{8^s} - \frac{1}{10^s} - \cdots. \quad (2.4.38)$$

Our goal is to evaluate the sum

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{5} + \frac{1}{7} - \frac{1}{8} + \frac{1}{10} - \cdots, \quad (2.4.39)$$

which correponds to $L(1, \chi_{-3})$ as we needed. We begin by recalling a result that we are going to use.

Theorem 2.4.10. *(Abel's Theorem)*

1. *Suppose that an infinite series converges, say* $\sum_{k=0}^{\infty} a_k$. *For* $0 \leq x \leq 1$, *form the power series* $f(x) = \sum_{k=0}^{\infty} a_k x^k$. *Then*

$$\sum_{k=0}^{\infty} a_k = \lim_{x \longrightarrow 1} f(x). \tag{2.4.40}$$

2. *Suppose that in an infinite series* $\sum_{k=0}^{\infty} a_k$, *not necessarily convergent, the partial sums* $\sum_{k=0}^{n} a_k$ *are all bounded in absolute value. Then the series* $\sum_{k=0}^{\infty} a_k k^{-\sigma}$ *converges for all* $\sigma > 0$.

Using the first part of Theorem 2.4.10, we define for $|x| < 1$ the Taylor series $f_{-3}(x)$ such that

$$\begin{aligned} f_{-3}(x) &= \sum_{n=1}^{\infty} \frac{\chi_{-3}(n)}{n} x^n, \\ &= x - \frac{x^2}{2} + \frac{x^4}{4} - \frac{x^5}{5} + \frac{x^7}{7} - \frac{x^8}{8} + \frac{x^{10}}{10} - \cdots . \end{aligned}$$

If we find an expression which has a closed form with $f_{-3}(x)$ that is continuous at $s = 1$, then what we need is in fact $f_{-3}(1)$. For that, we consider

$$f_{-3}(x) = \sum_{k=0}^{\infty} \left[\frac{x^{3k+1}}{3k+1} - \frac{x^{3k+2}}{3k+2} \right], \tag{2.4.41}$$

and this is possible as χ_{-3} is a character modulo 3. Also we have $f_{-3}(x) = L(1, \chi_{-3})$.

However, note that the derivative of f_{-3} is much easier to deal with rather than f_{-3} itself. So, we have

$$\begin{aligned} f'_{-3}(x) &= \sum_{k=0}^{\infty} \left(x^{3k} - x^{3k+1} \right), \\ &= (1-x) \sum_{k=0}^{\infty} x^{3k}, \\ &= \frac{1-x}{1-x^3}, \\ &= \frac{1}{1+x+x^2}. \end{aligned}$$

The point is that we would like to have a simple expression of $f_{-3}(x)$. Moreover, note that

$$f'_{-3}(x) = \frac{1}{1+x+x^2} = \frac{1}{(x+\frac{1}{2})^2 + \frac{3}{4}}. \tag{2.4.42}$$

Now, in order to get $f_{-3}(x)$ we just integrate (2.4.42). Since $f_{-3}(0) = 0$, we have

$$\begin{aligned} f_{-3}(x) &= \int_0^x \frac{dt}{\left(t+\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2}, \\ &= \frac{2}{\sqrt{3}} \arctan\left(\frac{2x+1}{\sqrt{3}}\right) - \frac{\pi}{3\sqrt{3}}. \end{aligned}$$

Therefore, plugging $x = 1$, we finally obtain

$$\begin{aligned} L(1,\chi_{-3}) &= f_{-3}(1), \\ &= \frac{2}{\sqrt{3}} \arctan\left(\frac{3}{\sqrt{3}}\right) - \frac{\pi}{3\sqrt{3}}, \\ &= \frac{2}{\sqrt{3}}\left(\frac{\pi}{3}\right) - \frac{\pi}{3\sqrt{3}}, \end{aligned}$$

that is,

$$L(1,\chi_{-3}) = \frac{\pi}{3\sqrt{3}}. \tag{2.4.43}$$

As one can see, Equation (2.4.43) is a particular case of Theorem 2.4.1. We have that, for $\Delta = -3$,

$$L(1,\chi_{-3}) = \frac{\pi}{3\sqrt{3}} = \frac{2\pi h(-3)}{\omega\sqrt{3}}, \tag{2.4.44}$$

which provides $h(-3) = 1$ because $\omega = 6$. The fact that $h(\Delta) = 1$, for $\Delta = -3$, is obtained from Theorem 2.1.10. To see this, by Condition (2.1.10), we get

$$|b| \leq \sqrt{\frac{|\Delta|}{3}} \quad \text{and} \quad b \equiv \Delta \bmod 2. \tag{2.4.45}$$

Moreover, we also have

$$4ac = b^2 - \Delta. \tag{2.4.46}$$

Using Equation (2.4.45), when $\Delta = -3$, we deduce that $b = 1$. By Equation (2.4.46), the only possible choice of a should divide $\frac{1}{4}(b^2 - \Delta) = 1$. Hence, $a = 1$. It follows that

$$c = \frac{b^2 - \Delta}{4a} = 1. \tag{2.4.47}$$

Thus, we get only one reduced form which is $x^2 + xy + y^2$, that is $h(-3) = 1$.

Chapter 3

A torsion bound for CM Elliptic Curves

In this chapter, we would like to highlight a paper of P. Clark and P. Pollack [4], in which they used the Siegel zero bound $1-\beta \geq C/\sqrt{|\Delta|}$ for imaginary quadratic fields to show a universal upper bound on the size of the torsion subgroup of a CM elliptic curve. This is one of our motivations to make the Siegel zero bound explicit.

3.1 Elliptic Curves

This section aims to make a brief introduction of elliptic curves. Most of the results here are classic and we assume that the reader is familiar with the basic notions of algebraic geometry.

3.1.1 Cubic Curves

The degree of a curve is determined by the degree of its defining polynomial. In particular, we are concerned with a curve of degree 3.

Definition 3.1.1. We define a *cubic curve* to be the curve defined by the equation

$$ax^3 + by^3 + cx^2y + dxy^2 + ex^2 + fy^2 + gxy + hx + iy + j = 0, \qquad (3.1.1)$$

where all the coefficients belong to some field.

By homogenization, the polynomial involved in (3.1.1) becomes homogeneous of degree 3. Such polynomials with some additional conditions have a particular structures.

Definition 3.1.2. Let K be a field. We define an *elliptic curve* over K, often denoted by $E_{/K}$, to be a nonsingular projective cubic curve together with a distinguished point that we denote by O.

If we want to consider points on an elliptic curve $E_{/K}$ with coordinates in some field $L \supset K$, we write $E(L)$.

We should also note that one can find another way of defining an elliptic curve. In particular, one of them can be derived from the following.

Definition 3.1.3. We define a *general Weierstrass equation* to be the equation given by

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3, \tag{3.1.2}$$

where $a_i \in K$ for all i.

An elementary property of a general Weierstrass equation is given as follows (see [7, p.468]).

Proposition 3.1.4. *Let C be a curve over a field K defined by a general Weierstrass equation $f(X, Y, Z) = 0$, where*

$$f(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - (X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3).$$

Then C is an absolutely irreducible curve, that is, it is irreducible over any algebraic closure of K.

With some change of variables, and by assuming some conditions on the characteristic of the field (here, different from 2 and 3), we obtain a much simple form of Equation (3.1.2).

Definition 3.1.5. An equation of the form

$$Y^2Z = X^3 + aXZ + bZ^3 \tag{3.1.3}$$

is called a *simple Weierstrass equation*.

The *discriminant* of an elliptic curve given as in Equation 3.1.3 is defined to be the quantity

$$D := -16(4a^3 + 27b^2). \tag{3.1.4}$$

Moreover, the *j-invariant* of an elliptic curve, say E, is defined by

$$j(E) := 1728\frac{4a^3}{4a^3 + 27b^2}. \tag{3.1.5}$$

Let us see how all of these definitions are related. Indeed, every nonsingular cubic which has a rational point over a field K is isomorphic to a Weierstrass equation as in (3.1.3), at least when K does not have characteristic 2 or 3. In the case we would like to include the case of characteristic 2 and 3, we have to consider a more general equation as we give in Equation (3.1.2). Here, we will not go through this process but for any further details, one may check for example [7, p.477].

Remark 5. In general, an elliptic curve can be defined by many forms other than by a Weierstrass equation. However, the most concrete description of an elliptic curve is obtained by the simple affine Weierstrass equation, written as

$$y^2 = x^3 + ax + b, \tag{3.1.6}$$

such that $4a^3 + 27b^2 \neq 0$. Notice that in the above situation, the characteristic of the field is assumed to be different from 2 and 3.

We note that the curve given by (3.1.3) has a point $O = [0 : 1 : 0]$, which is considered as the point at infinity on the curve (more precisely, when $Z = 0$). Moreover, the nonsingularity of the curve can be translated into the nonvanishing of its discriminant.

3.1.2 Group Law on a Cubic

Let C be a nonsingular cubic. Let P and Q be disctinct points on C. The ligne through P and Q cuts the cubic C at a third point, say R. Note that the point R may possibly equal to P or Q. In the case where $P = Q$, we operate similarly with the tangent line to the curve C at the point P.

Definition 3.1.6. Let us denote the point R described as above by $R = P \circ Q$. We define the law $+$ by choosing a point $O \in C$ called *origin* and setting $O' = O \circ O$, then

$$P + Q := O \circ (P \circ Q) \quad \text{and} \quad -P := O' \circ P. \tag{3.1.7}$$

This procedure of defining the law $+$ is known as the *chord-tangent method*. We recall that $C(K)$ denotes the set of rational points on K, that is, $C(K) = \{[X : Y : Z] \mid F(X, Y, Z) = 0\}$ for a degree 3 homogeneous polynomial $F \in K[X, Y, Z]$. A classical result about this is the following.

Theorem 3.1.7. *If $O \in C(K)$, then $C(K)$ has a structure of an abelian group.*

The proof can be found in [15, p.170]. Note that the construction used in the proof makes use of Bézout's theorem

From now on, we consider the law "+" with a curve defined by the general Weierstrass equation as defined in (3.1.2) and also $O = [0, 1, 0]$.

Proposition 3.1.8. *Let E be an elliptic curve defined by the (general) affine Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$. Then the following holds:*

- *Let $P = (x_1, y_1) \in E(K)$. Then $-P = (x, -y - a_1x_1 - a_3)$.*
- *Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$. Then*

– *if* $x_1 = x_2$ *and* $y_1 = -y_2 - a_1x_2 - a_3$, *then* $P_1 + P_2 = O$.
– *otherwise, let* $P_1 + P_2 = P_3 = (x_3, y_3)$. *Then*

$$\begin{cases} x_3 &= \lambda^2 + a_1\lambda - a_2 - x_1 - x_2, \\ y_3 &= -(\lambda + a_1)x_3 - \mu - a_3, \end{cases} \tag{3.1.8}$$

where $\lambda, \mu \in K$ *are given as follows:*

* *if* $x_1 \neq x_2$,

$$\begin{cases} \lambda &= \frac{y_2 - y_1}{x_2 - x_1}, \\ \mu &= \frac{y_1x_2 - y_2x_1}{x_2 - x_1}. \end{cases} \tag{3.1.9}$$

* *if* $x_1 = x_2$,

$$\begin{cases} \lambda &= \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3}, \\ \mu &= \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}. \end{cases} \tag{3.1.10}$$

One can find a proof of this proposition in [7, Section 7, Part II] or in Chapter III of [22].

Remarks 2. 1. By writing the Weierstrass equation as $f(x, y) = 0$, we notice that the case $P_1 = P_2$ gives us $\lambda = -\left(\frac{\partial f}{\partial x}\right)\Big/\left(\frac{\partial f}{\partial y}\right)$, as it must be.

2. A particular case of Proposition 3.1.8 is when $P_1 \neq \pm P_2$, so that

$$x(P_1 + P_2) = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 + a_1\left(\frac{y_2 - y_1}{x_2 - x_1}\right) - a_2 - x_1 - x_2. \tag{3.1.11}$$

In addition, we have also the so called *duplication formula* for $P = (x, y) \in E(K)$ such that

$$x(2P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4 + b_6}, \tag{3.1.12}$$

where the b_i is given as

$$\begin{aligned} b_2 &= a_1^2 + 4a_4, \\ b_4 &= 2a_4 + a_1a_3, \\ b_6 &= a_3^2 + 4a_6, \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2. \end{aligned}$$

3. For the case of the simple Weierstrass equation of the form $y^2 = x^3 + ax + b$, the above proposition is reduced as follows

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{if } P_1 \neq P_2, \\ \frac{3x_1^2 + a}{2y_1} & \text{if } P_1 = P_2, \end{cases} \tag{3.1.13}$$

and then

$$P_3 = (x_3, y_3) = (\lambda^2 - x_1 - x_2, \lambda(x_1 - x_3) - y_1). \tag{3.1.14}$$

Now, we state an important and remarkable result which concerns the group of rational points on an elliptic curve.

Theorem 3.1.9 (Mordell-Weil)**.** *[22, Theorem 6.7] Let E be an elliptic curve over a number field K. Then the group $E(K)$ is finitely generated.*

In particular, the case where $K = \mathbb{Q}$ is known as *Mordell's Theorem* (see [22, Theorem 4.1]).

Note that from the previous theorem together with Theorem 3.1.7, $E(K)$ is a finitely generated abelian group. Hence, the group $E(K)$, which is known as *Mordell-Weil group*, can be written in the form

$$E(K) \cong \mathbb{Z}^r \times E(K)[tors], \tag{3.1.15}$$

where the integer r is called the *rank* of the elliptic curve and $E(K)[tors]$ is the torsion subgroup.

We recall the next result in which one can derive Equation 3.1.15. In fact, it is based on the structure theorem of a finitely generated abelian group.

Proposition 3.1.10. *Let G be a finitely generated abelian group and T be the torsion subgroup of G. Then the quotient G/T is free. More precisely, we have $G = T \oplus F$ such that F is free and $F \cong G/T$, and also*

$$T = (a_1) \oplus \cdots \oplus (a_r)$$

where $o(a_i) = \varepsilon_i$ and $1 < \varepsilon_1 \mid \varepsilon_2 \mid \cdots \mid \varepsilon_r$.

3.2 Torsion Points on Elliptic Curves

A point of finite order is usually called *torsion point*. In this section, we discuss about this kind of point such as 2-torsion and 3-torsion points. Specifically, we allow the coordinates of a point on an elliptic curve to be complex.

Definition 3.2.1. Let $E_{/K}$ be an elliptic curve and let $n \geq 1$ be integer. The *n-torsion subgroup of E*, usually denoted by $E[n]$, is the set of points of order n.

More precisely, we have

$$E[n] = \{P \in E(\overline{K}) \mid nP = O\}. \tag{3.2.1}$$

Hence,

$$E(\overline{K})[tors] = \bigcup_{n=1}^{\infty} E[n]. \tag{3.2.2}$$

Remark 6. The previous notation nP means precisely

$$nP = \underbrace{P + \cdots + P}_{n \text{ summands}}. \tag{3.2.3}$$

Now, we are going to see two situations: points of order 2 and points of order 3.

We first begin with points of order 2, and let $P = (x, y) \in C(K)$ be such a point. By the definition, we have that $2P = O$ with $P \neq O$. In other words, this can be also seen as $P = -P$ and remember that we have $-P = (x, -y)$. Clearly, the previous equality (i.e. $P = -P$) holds only when $y = 0$. This means that the equation of the elliptic curve becomes $0 = f(x) = x^3 + ax^2 + bx + c$. Since the curve C is nonsingular, that is $f(x)$ and $f'(x)$ does not vanish simultaneously, then the polynomial $f(x)$ has three distinct roots. Thus, we obtain three distinct points, namely $P_1 = (x_1, 0)$, $P_2 = (x_2, 0)$ and $P_3 = (x_3, 0)$. Therefore, together with the point O, we have that

$$C[2] = \{O, P_1, P_2, P_2\} \tag{3.2.4}$$

We consider now the case of the points of order 3. Let $P = (x, y) \in C(K)$ be a point of order 3. In fact, this point satisfies $3P = O$ or $2P = -P$ with $P \neq O$. In other words, we obtain $x(2P) = x(-P)$ and recall that $x(P) = x(-P)$. Using (3.1.12), we have that

$$\frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c} = x, \tag{3.2.5}$$

$$x^4 - 2bx^2 - 8cx + b^2 - 4ac = 4x^4 + 4ax^3 + 4bx^2 + 4cx, \tag{3.2.6}$$

which implies $\varphi(x) = 3x^4 + 4ax^3 + 6bx^2 + 12cx - b^2 + 4ac = 0$. That is, points of order three should satisfy $\varphi(x) = 0$.

Now, we use the regular equation on adding two points in (3.1.10), and so

$$x(2P) = \left[\frac{f'(x)}{2y}\right]^2 - a - x - x, \tag{3.2.7}$$

$$= \frac{(f'(x))^2}{4f(x)} - a - 2x, \tag{3.2.8}$$

$$= x, \qquad \text{which also holds.} \tag{3.2.9}$$

From the last two equalities, we may write

$$2f(x)(6x + 2a) - (f'(x))^2 = 0. \tag{3.2.10}$$

However, note that we also have $f''(x) = 6x + 2a$. In fact, we get another way of writing $\varphi(x)$, such that

$$\varphi(x) = 2f(x)f''(x) - (f'(x))^2. \tag{3.2.11}$$

We claim that $\varphi(x)$ has four distinct roots. Indeed, it is enough to see if $\varphi(x)$ and its derivative have no common roots. We have

$$\begin{aligned} \varphi'(x) &= 2f'(x)f''(x) + 2f(x)f'''(x) - 2f'(x)f''(x), && (3.2.12)\\ &= 2f(x)f'''(x), && (3.2.13)\\ &= 12f(x), \quad \text{since } f'''(x) = 6. && (3.2.14) \end{aligned}$$

Assume now that we have a common root x_0, that is the two equations $\varphi(x_0) = 0$ and $\varphi'(x_0) = 0$ hold. In other words,

$$\begin{cases} 2f(x_0)f''(x_0) - (f'(x_0))^2 &= 0,\\ 12f(x_0) &= 0. \end{cases} \tag{3.2.15}$$

Thus, $f(x_0) = 0$ implies $f'(x_0) = 0$ and this contradicts the nonsingularity of the cubic curve. Therefore, $\varphi(x)$ has four distinct roots, namely x_i for $i = 1, \ldots, 4$. For each x_i, we have a pair $\pm\sqrt{f(x_i)}$ so that we obtain eight points of the form $P_i = (x_i, \pm\sqrt{f(x_i)})$ for $i = 1, \ldots, 4$. Including O, it follows that there are precisely nine points of order 3, that is

$$C[3] = \left\{O, (x_i, \pm\sqrt{f(x_i)})\right\} \quad \text{with } i = 1, 2, 3, 4. \tag{3.2.16}$$

One can notice from the above discussion that given an integer $n \neq 0$, there are exactly n^2 points of order n. In fact, this is true for any positive integer n as we see below.

Theorem 3.2.2. *[26, p.79] Let $E_{/K}$ be an elliptic curve and let n be a positive integer. If the characteristic of K does not divide n, or is 0, then we have*

$$E[n] \cong \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}. \tag{3.2.17}$$

A remarkable theorem about torsion points is also due to Lutz and Nagell. This result has a strong use in terms of computation. In particular, it can be used to check that a given point has an infinite order. We state *Nagell-Lutz's theorem.*

Theorem 3.2.3. *[23, p.56] Let E be an elliptic curve as in* (3.1.6) *such that a, b and c are integers. Let D be the discriminant of the cubic polynomial as in* (3.1.4)*. Let $P = (x, y)$ be a rational point of finite order. Then x and y are integers; and either $y = 0$, in which case P has order two, or else y^2 divides D.*

We have another result concerning the determination of which orders are possible for a point on a nonsingular rational cubic curve. The following is known as *Mazur's theorem.*

Theorem 3.2.4. *[22, p. 242] Let $E_{/\mathbb{Q}}$ be an elliptic curve. Then the torsion subgroup $E(\mathbb{Q})[tors]$ is isomorphic to one of the following group:*

$$\begin{cases} \mathbb{Z}/n\mathbb{Z} & \text{with } 1 \leq n \leq 10 \text{ or } n = 12, \\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z} & \text{with } 1 \leq n \leq 4. \end{cases} \tag{3.2.18}$$

Let us look now on a general case where K is any subfield of $\mathbb{C}$. Suppose that K is a Galois extension of $\mathbb{Q}$. Then for any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ and any point $P \in E(K)$, we can define a new point

$$\sigma(P) = (\sigma(x), \sigma(y)). \tag{3.2.19}$$

The following result provides informations on the point $\sigma(P)$.

Proposition 3.2.5. *[23, p.215] Let $E_{/\mathbb{Q}}$ be an elliptic curve and let K be a Galois extension of $\mathbb{Q}$.*

- *The set of points $E(K)$ is a subgroup of $E(\mathbb{C})$.*
- *For $P \in E(K)$ and $\sigma \in Gal(K/\mathbb{Q})$, define*

$$\sigma(P) = \begin{cases} (\sigma(x), \sigma(y)) & \text{if } P = (x, y), \\ O & \text{if } P = O. \end{cases}$$

  Then $\sigma(P) \in E(K)$.
- *For all $P \in E(K)$ and all $\sigma, \tau \in Gal(K/\mathbb{Q})$, we have $(\sigma\tau)(P) = \sigma(\tau(P))$. Furthermore, the identity element $\iota \in Gal(K/\mathbb{Q})$ acts trivially, i.e. $\iota(P) = P$.*
- *For all $P, Q \in E(K)$ and all $\sigma \in Gal(K/\mathbb{Q})$, we have $\sigma(P+Q) = \sigma(P) + \sigma(Q)$ and $\sigma(-P) = -\sigma(P)$. In particular, $\sigma(nP) = n(\sigma(P))$ for any integers n.*
- *Let $P \in E(K)$ be a point of finite order n and let $\sigma \in Gal(K/\mathbb{Q})$. Then $\sigma(P)$ also has order n.*

3.3 Complex Multiplication: basic idea and aspect

In this section, we are interested on an elliptic curve over $\mathbb{C}$. More specifically, we are going to develop the phenomen of complex multiplication. This is best understood with the help of an example. So, we put a basic example with our discussion.

Let n be any integer and $E_{/\mathbb{C}}$ be an elliptic curve. We have an endomorphism of $E(\mathbb{C})$ given as follows

$$[n]: \ E(\mathbb{C}) \to E(\mathbb{C}), \ P \longmapsto nP. \tag{3.3.1}$$

In fact, this is usually called *multiplication-by-n map*. One may quickly notice that the kernel of this homomorphism is the subgroup of n-torsion points $E[n] := \{P \in E(\mathbb{C}) \mid nP = O\}$.

Now, let us see an example. We consider the elliptic curve $E_{1/\mathbb{C}}$ defined by the equation $y^2 = x^3 + x$. Define the map ψ on $E_1(\mathbb{C})$ by

$$\psi: \ E_1(\mathbb{C}) \to E_1(\mathbb{C}), (x, y) \mapsto (-x, iy). \tag{3.3.2}$$

Clearly, we see that

$$\begin{aligned}(iy)^2 &= -y^2, \\ &= -(x^3 + x), \\ &= (-x)^3 + (-x).\end{aligned}$$

It follows that $(-x, iy)$ belongs to the group $E_1(\mathbb{C})$. Moreover, using the addition law of points in $E_1(\mathbb{C})$ we deduce that ψ is a homomorphism.

Definition 3.3.1. We define an *isogeny* to be a non-trivial homomorphism between two elliptic curves given by rational functions.

We are mainly concerned with homomorphism of elliptic curve to itself, that is an endomorphism of elliptic curve. Explicitly, given an elliptic curve $E_{/\mathbb{C}}$, we have a homomorphism ψ such that

$$\psi(x, y) = \Big(U(x, y), V(x, y)\Big), \tag{3.3.3}$$

where $U(x, y)$ and $V(x, y)$ are rational functions.

For an integer n, we know that every elliptic curve has an endomorphism given by the mutliplication-by-n map. However, some elliptic curves have an additional endomorphism. For example, we have seen that (3.3.2) defines a homomorphism on the curve E_1 given by $y^2 = x^3 + x$. Such an elliptic curve has in fact a special definition.

Definition 3.3.2. Let n be an integer. We say that an elliptic curve $E_{/K}$ has *complex mutliplication*, or for short of *CM-type*, if there is an endomorphism $\psi : E(K) \to E(K)$ that is not a multiplication-by-n map.

Following our previous example, we say that $E_{1/\mathbb{C}}$ has a complex multiplication given by (3.3.2).

More generally, since the set $E(K)$ has an abelian group structure one may consider its ring of endomorphisms which is often denoted by $\mathrm{End}(E)$. Note

that the addition is induced by that of $E(K)$ and the multiplication is given by the composition of endomorphisms.

Let us describe the general situation about the phenomenon of complex multiplication. We consider an elliptic curve E over $\mathbb{C}$ or more generally over any field of characteristic zero. We know that the ring $\mathrm{End}(E)$ of endomorphisms of an elliptic curve E contains the maps $[n]$ for any $n \in \mathbb{Z}$. The main point is that $\mathrm{End}(E)$ is equal to either

1. $\mathbb{Z}$ that is there are no additional endomorphisms, or
2. an order in an imaginary quadratic field, in which case $\mathrm{End}(E) \cong \mathbb{Z} + \mathbb{Z}\tau$ such that
$$\tau = \frac{\Delta + \sqrt{\Delta}}{2}, \tag{3.3.4}$$
for some $\Delta < 0$ and $\Delta \equiv 0$ or $1 \bmod 4$.

More precisely, in the second case we say that E has complex multplication by the order of discriminant Δ. Note also that the concerned order is not necessarily maximal.

Now, let us see how is the situation over $\mathbb{Q}$. The theory of complex multiplications over $\mathbb{Q}$ provides a remarkable result, in particular the description of an abelian extension of an imaginary quadratic field. In fact, this is mainly based on the knowledge of the field obtained by adding the coordinates of all torsion points of an elliptic curve to some specific field.

Let n be an integer. Recall that a point $P \in E[n]$ has two coordinates. We are concerned with the field generated by all of the coordinates of all the n-torsion points of E.

We have the following result:

Proposition 3.3.3. *[23, p.218] Let $E_{/\mathbb{Q}}$ be an elliptic curve given by a Weierstrass equation*
$$y^2 = x^3 + ax^2 + bx + c.$$

1. *Let $P = (x_1, y_1) \in E[n]$ be a point of order n. Then x_1 and y_1 are algebraic over $\mathbb{Q}$.*
2. *Let $E[n] = \{(x_1, y_1), \cdots, (x_m, y_m), O\}$ be the complete set of points of $E(\mathbb{C})$ of order n where $m = n^2 - 1$. Let*
$$K = \mathbb{Q}(x_1, y_1, \ldots, x_m, y_m) \tag{3.3.5}$$
be the field generated by the coordinates of all the points in $E[n]$. Then K is a Galois extension of $\mathbb{Q}$. In general, $Gal(K/\mathbb{Q})$ will not be abelian.

According to this proposition, one may notice that adjoining the coordinates of torsion points to $\mathbb{Q}$ does not give necessarily a Galois extension which

is abelian. In the opposite direction, let us treat a (basic) case in which we obtain an abelian extension. For that, we shall consider an elliptic curve of CM-type.

DETERMINING THE ABELIAN EXTENSION OF $\mathbb{Q}(i)$:

In what follows, we consider the elliptic curve E_1 defined by the equation $y^2 = x^3 + x$, and from now on by E we mean the curve E_1. We saw previously in Equation (3.3.2) that this curve has complex multiplication given by

$$\begin{aligned} \varphi \ : E(\mathbb{C}) &\longrightarrow E(\mathbb{C}), \\ (x, y) &\mapsto (-x, iy). \end{aligned}$$

Let K be a Galois extension of $\mathbb{Q}$ such that $i \in K$. Let $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$. Then for any point $P \in E(K)$, we saw from Proposition 3.2.5 that $\sigma(P)$ and $\varphi(P)$ are respectively points in $E(K)$. We are now interested in the following situation: for any point $P \in E(K)$ we ask if

$$\sigma(\varphi(P)) = \varphi(\sigma(P)) \tag{3.3.6}$$

holds. Explicitly, we have

$$\begin{aligned} \sigma(\varphi(P)) &= \sigma(-x, iy) = (-\sigma(x), \sigma(i)\sigma(y)), \\ \varphi(\sigma(P)) &= \varphi(\sigma(x), \sigma(y)) = (-\sigma(x), i\sigma(y)). \end{aligned}$$

Clearly, one can see that the commutativity of σ and φ holds if we have

$$\sigma(i) = i. \tag{3.3.7}$$

Thus, we now switch to the Galois extension of $\mathbb{Q}(i)$ instead of considering $\mathbb{Q}$. This is because Condition (3.3.7) is realized if we take $\sigma \in \mathrm{Gal}(K/\mathbb{Q}(i))$ instead of $\mathrm{Gal}(K/\mathbb{Q})$.

We recall the following result on Galois representation which can be found in [23, p.225].

Theorem 3.3.4. *Let E be an elliptic curve given by a Weierstrass equation with rational coefficients, and let $n \geq 2$ be an integer. Fix generators P_1 and P_2 for $E[n]$. Then the map*

$$\rho_n \ \ Gal(\mathbb{Q}(E[n])/\mathbb{Q}) \longrightarrow GL_2(\mathbb{Z}/n\mathbb{Z}), \tag{3.3.8}$$

is a one-to-one group homomorphism.

In our case, note that the points of order two must satisfy $x^3 + x = 0$. More precisely, we have

$$E[2] = \{(0, 0), (i, 0), (-i, 0), O\}.$$

Hence, we obtain $\mathbb{Q}(E[2]) = \mathbb{Q}(i)$. Moreover, the extension $\mathbb{Q}(E[2])/\mathbb{Q}$ is Galois since $\mathbb{Q}(i)$ is the splitting field of the minimum polynomial x^2+1. So it follows that $\mathrm{Gal}(\mathbb{Q}(E[2])/\mathbb{Q}) = \{\sigma, \tau\}$ such that σ is the identity map (i.e. $i \mapsto i$) and τ is the complex conjugation (i.e. $i \mapsto -i$). Take $P_1 = (i, 0)$ and $P_2 = (-i, 0)$ as the generators for $E[2]$. Then we get

$$\begin{aligned}\tau(P_1) &= \tau(i,0) = (-i,0),\\ \tau(P_2) &= \tau(-i,0) = (i,0).\end{aligned}$$

So, for $n = 2$, Theorem 3.3.4 tells us that

$$\rho_2(\sigma) = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \text{ and } \rho_2(\tau) = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}.$$

More generally, by Theorem 3.3.4, we may have

$$\begin{aligned}\rho_n \;:\; \mathrm{Gal}(\mathbb{Q}(i)(E[n])/\mathbb{Q}(i)) &\longrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z}),\\ \sigma &\longmapsto \left(\begin{smallmatrix}\alpha_\sigma & \beta_\sigma\\ \gamma_\sigma & \delta_\sigma\end{smallmatrix}\right)\end{aligned}$$

such that

$$\begin{aligned}\sigma(P_1) &= \alpha_\sigma P_1 + \gamma_\sigma P_2\\ \sigma(P_2) &= \beta_\sigma P_1 + \delta_\sigma P_2.\end{aligned}$$

On the other hand, let $\varphi : E[n] \longrightarrow E[n]$ be a homomorphism. Note that this is because, for any $P \in E[n]$, we have $n\varphi(P) = \varphi(nP) = \varphi(O) = O$. Using Theorem 3.3.4, there exist $a, b, c, d \in \mathbb{Z}/n\mathbb{Z}$ with

$$\begin{aligned}\varphi(P_1) &= aP_1 + cP_2\\ \varphi(P_2) &= bP_1 + dP_2.\end{aligned}$$

Also notice that (3.3.6) now holds as we consider $\sigma \in \mathrm{Gal}(\mathbb{Q}(i)(E[n])/\mathbb{Q}(i))$. In matrix form, it follows that

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\alpha_\sigma & \beta_\sigma\\ \gamma_\sigma & \delta_\sigma\end{bmatrix} = \begin{bmatrix}\alpha_\sigma & \beta_\sigma\\ \gamma_\sigma & \delta_\sigma\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}. \tag{3.3.9}$$

Now by using Lemma 6.20 and Lemma 6.21 from [23], we are able to state the following:

Theorem 3.3.5. *[23, p.236] Let E be the elliptic curve defined by the equation $y^2 = x^3 + x$. For any integer $n \geq 1$, let*

$$K_n = \mathbb{Q}(i)(E[n])$$

be the field generated by i and the coordinates of the points in $E[n]$. Then K_n is a Galois extension of $\mathbb{Q}(i)$, and its Galois group is abelian.

3.4 Truth about Torsion in the CM case

This section is based on the paper [4] of P.L. Clark and P. Pollack. Here, we are going to study this article and also to exhibit a specific step in the proof of a particular theorem.

3.4.1 Main statement

In 1973, A. Ogg presented the *torsion conjecture* or also known as *uniform boundedness conjecture*. For the case of elliptic curves over the rational field, it was solved by B. Mazur in his paper of 1977 and 1978 (see Theorem 3.2.4). The case of quadratic fields and number fields of degree at most 8 has been proved respectively by S. Kamienny (1992) and S. Kamienny & B. Mazur (1995). Then, L. Merel in 1996 showed the general case of elliptic curves over any number field. In his work of 1999, P. Parent gave an effective bound for the size of the torsion subgroup in terms of the degree of the number field.

In [4], the authors showed that:

Theorem 3.4.1. *There is an absolute, effective constant C such that for all number fields F of degree $d \geq 3$ and all elliptic curves E over F with complex multiplication, we have*

$$\#E(F)[tors] \leq Cd \operatorname{loglog} d. \tag{3.4.1}$$

In other words, this theorem says that the upper order of the size of the torsion subgroup of a CM elliptic curve over a degree d number field is d loglog d.

Aligned with this, M. Hindry and J. Silverman also proved in [16] that:

Theorem 3.4.2. *For all number fields F of degree $d \geq 2$ and all elliptic curves E over F with j-invariant $j(E) \in O_F$, we have*

$$\#E(F)[tors] \leq 1977408d \log d. \tag{3.4.2}$$

The above two results provide us informations on the size of the torsion subgroup of all elliptic curves over a number field of degree greater than 2. We should note that every CM elliptic curve E over F has $j(E) \in O_F$ but there are only finitely many $j \in O_F$ which are j-invariants of CM elliptic curve E over F. Hence, Theorem 3.4.1 has a strong hypothesis and a slightly different conclusion compared with Theorem 3.4.2.

Moreover, we also have the following result which is due to F. Breuer in [3].

Theorem 3.4.3. *Let E be an elliptic curve over a number field F. There exists a constant $c(E,F) > 0$, integers $3 \leq d_1 < d_2 \cdots < d_n < \cdots$ and number fields $F_n \supset F$ with $[F_n : F] = d_n$ such that for all $n \in \mathbb{Z}^+$ we have*

$$\#E(F)[tors] \geq \begin{cases} c(E,F)d_n \log\log d_n & \text{if } E \text{ has CM,} \\ c(E,F)\sqrt{d_n \log\log d_n} & \text{otherwise.} \end{cases} \tag{3.4.3}$$

Hence, if we use $T_{CM}(d)$ to denote the largest size of the torsion sugbroup of a CM elliptic curve over a degree d number field, $d \log\log d$ is its upper order, that is

$$\limsup_d \frac{T_{CM}(d)}{d \log\log d} \in]0, \infty[. \tag{3.4.4}$$

In fact, a more precise result is known. The following can be found in [5, Theorem 1.1].

Theorem 3.4.4. *We have*

$$\limsup_d \frac{T_{CM}(d)}{d \log\log d} = \frac{e^\gamma \pi}{\sqrt{3}}. \tag{3.4.5}$$

Note that γ in the previous theorem means the Euler-Mascheroni constant. That is the constant on the right hand side is 3.2305...

The main result in [4] is Theorem 3.4.1 and the proof of this was mainly based on three other results namely Theorem 5, Theorem 7 and Theorem 8 as in the paper. Here, we are going to state these theorems and particularly we will explore Theorem 8. The reason for that will become clear later.

3.4.2 Notions from class field theory

We summarize here most of the notions we will need to explore Theorem 3.4.1. We denote by: K an arbitrary number field, $h(O_K) = h_K$ the class number of K, $\omega = \omega_K$ the number of roots unity in K and $v_\mathfrak{p}$ the $\mathfrak{p}$-adic valuation.

We begin with some notions of class field theory. The basic idea of this theory is that the generalized ideal class group are the Galois group of all abelian extensions of K established by the Artin map. To make this precise, let us define these terms.

We recall that prime ideals of O_K are called *finite primes* to distinguish them from the *infinite primes*, which are determined by the embedding of K into $\mathbb{C}$. A *real* infinite prime is an embedding $\sigma : K \to \mathbb{R}$ and a *complex* infinite prime is a pair of complex conjugate embeddings $\sigma, \overline{\sigma} : K \to \mathbb{C}$ $(\sigma \neq \overline{\sigma})$. For an extension L of K, an infinite prime σ of K *ramifies* in L provided that σ is real but it has an extension to L which is complex. An extension L of K is *unramified* if it is unramified at all finite or infinite primes.

Definition 3.4.5. A *modulus* $\mathfrak{m}$ in K is a formal product

$$\mathfrak{m} = \prod_\mathfrak{p} \mathfrak{p}^{n_\mathfrak{p}}$$

over all primes (finite or infinite) $\mathfrak{p}$ of K such that $n_\mathfrak{p}$ must satisfy:

1. $n_\mathfrak{p} \leq 0$, and at most finitely many are nonzero,

2. $n_{\mathfrak{p}} = 0$ whenever $\mathfrak{p}$ is a complex infinite prime,

3. $n_{\mathfrak{p}} \geq 1$ whenever $\mathfrak{p}$ is a real infinite prime.

Thus, a modulus $\mathfrak{m}$ can be written as $\mathfrak{m} = \mathfrak{m}_0\,\mathfrak{m}_\infty$, where

$$\begin{cases} \mathfrak{m}_0 & \text{is an ideal of } O_K, \\ \mathfrak{m}_\infty & \text{is a product of distinct real infinite primes of } K. \end{cases}$$

In the case that all the exponents $n_{\mathfrak{p}}$ are zero, we set $\mathfrak{m} = 1$. For a modulus $\mathfrak{m}$, we use $\mathcal{I}_K(\mathfrak{m})$ to denote the group of all fractional ideals of O_K which are relatively prime to $\mathfrak{m}$ (i.e. relatively prime to $\mathfrak{m}_0$), and $\mathcal{P}_{K,1}(\mathfrak{m})$ to denote the subgroup of $\mathcal{I}_K(\mathfrak{m})$ generated by the principal ideals αO_K, where $\alpha \in O_K$ satisfies

$$\alpha \equiv 1 \bmod \mathfrak{m}_0 \text{ and } \sigma(\alpha) > 0 \text{ for any real infinite prime } \sigma \text{ dividing } \mathfrak{m}_\infty\,.$$

The quotient $\mathcal{I}_K(\mathfrak{m})/\mathcal{P}_{K,1}(\mathfrak{m})$ is called *ray class group* for $\mathfrak{m}$.

Definition 3.4.6. A subgroup $\mathcal{H} \subset \mathcal{I}_K(\mathfrak{m})$ is called a *congruence subgroup* for the modulus $\mathfrak{m}$ if it satisfies

$$\mathcal{P}_{K,1}(\mathfrak{m}) \subset \mathcal{H} \subset \mathcal{I}_K(\mathfrak{m}), \tag{3.4.6}$$

and the quotient

$$\mathcal{I}_K(\mathfrak{m})/\mathcal{H} \tag{3.4.7}$$

is called a *generalized ideal class group* for $\mathfrak{m}$.

We now aim to define the Artin symbol. To do so, we first recall the next lemma.

Lemma 3.4.7. *[8, p.95] Let $K \subset L$ be a Galois extension and $\mathfrak{p}$ be a prime of O_K which is unramified in L. If $\mathfrak{P}$ is a prime of O_L containing $\mathfrak{p}$, then there is a unique element $\sigma \in Gal(L/K)$ such that for all $\alpha \in O_L$*

$$\sigma(\alpha) \equiv \alpha^{N(\mathfrak{p})} \bmod \mathfrak{P}, \tag{3.4.8}$$

where $N(\mathfrak{p})$ is the norm of $\mathfrak{p}$.

Definition 3.4.8. The *Artin symbol* is the unique element $\sigma \in \mathrm{Gal}(L/K)$ defined in Lemma 3.4.7. Since it depends on the prime $\mathfrak{P}$ of L, we often denote it by $\left(\frac{L/K}{\mathfrak{P}}\right)$.

An important property of the Artin symbol is that for any $\alpha \in O_L$, we have

$$\left(\frac{L/K}{\mathfrak{P}}\right)(\alpha) \equiv \alpha^{N(\mathfrak{p})} \bmod \mathfrak{P},$$

where $\mathfrak{p} = \mathfrak{P} \cap O_K$. Moreover, we also have the following properties:

Proposition 3.4.9. *[8, p.96] Let $K \subset L$ be a Galois extension and $\mathfrak{p}$ be an unramified prime of K. Given a prime $\mathfrak{P}$ of L containing $\mathfrak{p}$, we have:*

1. *if $\sigma \in Gal(L/K)$, then*

$$\left(\frac{L/K}{\sigma(\mathfrak{P})}\right) = \sigma\left(\frac{L/K}{\mathfrak{P}}\right)\sigma^{-1}, \tag{3.4.9}$$

2. *the order of $\left(\frac{L/K}{\mathfrak{P}}\right)$ is the inertial degree $f := [O_L/\mathfrak{P} : O_K/\mathfrak{p}]$,*

3. *$\mathfrak{p}$ splits completely in L if and only if $\left(\frac{L/K}{\mathfrak{P}}\right) = 1$.*

In the case that the extension L/K is abelian, the Artin symbol $\left(\frac{L/K}{\mathfrak{P}}\right)$ depends only on the underlying prime $\mathfrak{p}$ such that $\mathfrak{p} = \mathfrak{P} \cap O_K$. Indeed, if $\mathfrak{P}'$ is another prime containing $\mathfrak{p}$, we know that $\sigma(\mathfrak{P}) = \mathfrak{P}'$ for some $\sigma \in$ Gal(L/K). From Proposition 3.4.9, it follows that

$$\left(\frac{L/K}{\mathfrak{P}'}\right) = \left(\frac{L/K}{\sigma(\mathfrak{P})}\right) = \sigma\left(\frac{L/K}{\mathfrak{P}}\right)\sigma^{-1} = \left(\frac{L/K}{\mathfrak{P}}\right)$$

and the last equality holds since $\mathrm{Gal}(L/K)$ is abelian. Hence, whenever the extension L/K is abelian the Artin symbol can also be written as $\left(\frac{L/K}{\mathfrak{p}}\right)$. Moreover, the situation is interesting when the abelian extension L/K is unramified. This is because the Artin symbol $\left(\frac{L/K}{\mathfrak{p}}\right)$ is defined for all primes $\mathfrak{p}$ of O_K. We are now ready to define the Artin map.

Definition 3.4.10. Let $\mathfrak{m}$ be a modulus divisible by all ramified primes of an abelian extension $K \subset L$. The *Artin map* for L/K and $\mathfrak{m}$ is a surjective group homomorphism

$$\Phi_{L/K,\mathfrak{m}} : \mathcal{I}_K(\mathfrak{m}) \longrightarrow \mathrm{Gal}(L/K) \tag{3.4.10}$$

such that, for any $\mathfrak{a} \in \mathcal{I}_K(\mathfrak{m})$ we have

$$\Phi_{L/K,\mathfrak{m}}(\mathfrak{a}) = \prod_{i=1}^{s}\left(\frac{L/K}{\mathfrak{p}_i}\right)^{r_i}, \tag{3.4.11}$$

where $\left(\frac{L/K}{\cdot}\right)$ is the Artin symbol (that is each $\mathfrak{p}_i$ does not divide $\mathfrak{m}$) and the factorization of $\mathfrak{a} = \prod_{i=1}^{s}\mathfrak{p}_i^{r_i}$ $(r_i \in \mathbb{Z})$ is given by Proposition 2.3.13.

We recall the next theorem which is known as the *Existence Theorem.*

Theorem 3.4.11. *[8, Theorem 8.6] Let $\mathfrak{m}$ be a modulus of K, and let $\mathcal{H}$ be a congruence subgroup for $\mathfrak{m}$. Then there is a unique abelian extension L of K, all of whose ramified primes divide $\mathfrak{m}$ such that if we have the Artin map as in* (3.4.10)*, then*

$$\mathcal{H} = Ker(\Phi_{L/K,\mathfrak{m}}). \tag{3.4.12}$$

Definition 3.4.12. The *ray class field* for the modulus $\mathfrak{m}$ is the unique abelian extension $K_{\mathfrak{m}}$ of K such that

$$\mathcal{P}_{K,1}(\mathfrak{m}) = \mathrm{Ker}(\Phi_{K_{\mathfrak{m}}/K,\mathfrak{m}}), \tag{3.4.13}$$

obtained from the Existence Theorem.

Now, for a nonzero ideal $\mathfrak{a}$ of O_K, we define

$$\varphi_K(\mathfrak{a}) = N(\mathfrak{a}) \prod_{\mathfrak{p}|\mathfrak{a}} \left(1 - \frac{1}{N(\mathfrak{p})}\right), \tag{3.4.14}$$

and also by $K^{(\mathfrak{a})}$ we mean the ray class field of K for the ideal $\mathfrak{a}$. The next lemma can be found in [6]. We have:

Lemma 3.4.13. *Let K be an imaginary quadratic field and $\mathfrak{a}$ an ideal of O_K. Then*

$$\frac{h(O_K)\varphi_K(\mathfrak{a})}{6} \leq \frac{h(O_K)\varphi_K(\mathfrak{a})}{\omega_K} \leq [K^{(\mathfrak{a})} : K] \leq h(O_K)\varphi_K(\mathfrak{a}). \tag{3.4.15}$$

We recall from our discussion in (3.3.4) that an elliptic curve E over a field of characteristic 0 has complex multiplication if $\mathbb{Z} \subsetneq \mathrm{End}(E)$. Hence, $\mathrm{End}(E)$ is an order in an imaginary quadratic field (see[8, Chap.2,§7]). In this case, we say that E has

$$\begin{cases} O_K\text{-CM} & \text{if } \mathrm{End}(E) \cong O_K, \\ K\text{-CM} & \text{if } \mathrm{End}(E) \text{ is an order in } K. \end{cases} \tag{3.4.16}$$

We now state respectively the two results which are the building blocks of the proof of Theorem 3.4.2.

Theorem 3.4.14. *[4, Theorem 5] Let K be an imaginary quadratic field, $F \supset K$ a number field, E a K-CM elliptic curve over F and n a positive integer. If $(\mathbb{Z}/n\mathbb{Z})^2 \hookrightarrow E(F)$, then we have $F \supset K^{(nO_K)}$.*

The next theorem is somehow called as squaring the torsion subgroup of a CM elliptic curve.

Theorem 3.4.15. *[4, Theorem 7] Let K be an imaginary quadratic field, let $F \supset K$ a field extension, and let E be a K-CM elliptic curve over F. Assume that for positive integers a and b we have an injection $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/ab\mathbb{Z} \hookrightarrow E(F)$. Then $[F(E[ab]) : F] \leq b$.*

3.4.3 Proof of the torsion bound

Here, we adopt the following notations: let $f(x)$ and $g(x)$ be two functions defined for all sufficiently large x;

- $f(x) = O(g(x))$ means that there exists constants x_0 and c such that $|f(x)| \leq c|g(x)|$ for all $x \geq x_0$,
- $f(x) \ll g(x)$ is equivalent to $f(x) = O(g(x))$. It is also known as Vinogradov's notation,
- $f(x) \asymp (g(x))$ means that both $f(x) \ll g(x)$ and $g(x) \ll f(x)$ hold. In this case, we say that f and g have the same order of magnitude.

Now, we are concerned with the next theorem:

Theorem 3.4.16. *[4, Theorem 8] There is a positive effective constant* C *such that for all imaginary quadratic fields K and all nonzero ideals $\mathfrak{a}$ of O_K with $N(\mathfrak{a}) \geq 3$, we have*

$$\varphi_K(\mathfrak{a}) \geq \frac{\mathrm{C}}{h_K} \frac{N(\mathfrak{a})}{\operatorname{loglog} N(\mathfrak{a})}. \tag{3.4.17}$$

To prove this, we further need the following lemma.

Lemma 3.4.17. *[4, Lemma 9] There is an effective constant C, not the same as in* (3.4.17)*, for which the following holds: Let Δ be a fundamental discriminant with $\Delta < 0$, let $K = \mathbb{Q}(\sqrt{\Delta})$, and let $\chi(\cdot) = \chi_\Delta(\cdot)$ be the Kronecker symbol (see Equation* (2.2.8)*). For all $x \geq 2$,*

$$\prod_{p \leq x} \left(1 - \frac{\chi(p)}{p}\right) \geq \frac{C}{h_K}. \tag{3.4.18}$$

Proof. From the class number formula (see Theorem 2.4.1), we can write $L(1,\chi)\sqrt{|\Delta|} \asymp h_K$. For any $x \geq 2$, note that we may write

$$L(1,\chi) = \prod_{p} \left(1 - \frac{\chi(p)}{p}\right)^{-1}, \tag{3.4.19}$$

$$= \prod_{p \leq x} \left(1 - \frac{\chi(p)}{p}\right)^{-1} \prod_{p > x} \left(1 - \frac{\chi(p)}{p}\right)^{-1}. \tag{3.4.20}$$

Now with some rearrangement, we have that Equation (3.4.18) holds if and only if

$$\prod_{p > x} \left(1 - \frac{\chi(p)}{p}\right) \ll \sqrt{|\Delta|}. \tag{3.4.21}$$

Note that the product on the left-hand side of (3.4.21) indexed by $p \leq \exp(\sqrt{|\Delta|})$ makes a contribution of $O(\sqrt{|\Delta|})$. Indeed, this is because by taking logarithms, we get

$$\begin{aligned}\left|\log \prod_{p\leq \exp(\sqrt{|\Delta|})}\left(1-\frac{\chi(p)}{p}\right)\right| &= \left|\sum_{p\leq \exp(\sqrt{|\Delta|})} \log\left(1-\frac{\chi(p)}{p}\right)\right| \\ &\leq \sum_{p\leq \exp(\sqrt{|\Delta|})}\left|\frac{\chi(p)}{p}\right| \\ &\leq \sum_{p\leq \exp(\sqrt{|\Delta|})}\frac{1}{p} \\ &\leq \log\Big(\sqrt{|\Delta|}\Big)+K\end{aligned}$$

that is

$$\prod_{p\leq \exp(\sqrt{|\Delta|})}\left(1-\frac{\chi(p)}{p}\right) \leq e^K\sqrt{|\Delta|}.$$

Denote $y = \max\{x, \exp(\sqrt{|\Delta|})\}$. It suffices now to show that $\prod_{p>y}(1-\chi(p)/p)$ is bounded above by a constant. In other words, this holds if one could prove that $\sum_{p>y}\frac{\chi(p)}{p}$ is equal to $O(1)$, which is obtained by taking logarithms.

We assume that $L(s,\chi)$ has a Siegel zero β, otherwise the argument is the same but simpler. For $t \geq \exp(\sqrt{|\Delta|})$, we recall from [9, p.123] that the explicit formula gives

$$S(t) := \sum_{p\leq t}\chi(p)\log p = -\frac{t^\beta}{\beta}+O\left(\frac{t}{\log t}\right). \tag{3.4.22}$$

By partial summation, we have

$$\sum_{p>y}\frac{\chi(p)}{p} = -\frac{S(y)}{y\log y}+\int_y^\infty \frac{S(t)}{t^2(\log t)^2}(1+\log t)dt, \tag{3.4.23}$$

$$\ll 1+\int_y^\infty \frac{t^\beta}{t^2\log t}dt. \tag{3.4.24}$$

Using a bound on β from [13], that is $\beta \leq 1-\frac{C}{\sqrt{|\Delta|}}$, and since $y \geq \exp(\sqrt{|\Delta|})$, we obtain that the final integral is at most

$$\int_{\exp(\sqrt{|\Delta|})}^\infty \frac{\exp\Big(-\frac{C}{\sqrt{|\Delta|}}\log t\Big)}{t\log t}\,dt. \tag{3.4.25}$$

Denote $u = \log t/\sqrt{|\Delta|}$. Then, the above integral becomes $\int_1^\infty \exp(-Cu)u^{-1}\,du$ which is convergent, and this is as we needed. □

Now, let us prove the theorem concerning the torsion bound.

Proof of Theorem 3.4.2: Let F be a number field of degree $d \geq 3$ and let E be a K-CM elliptic curve over F. We may suppose that $\#E(F)[tors] \geq 3$. Using Proposition 3.1.10, we have $\#E(FK)[tors] \cong \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/ab\mathbb{Z}$ for positive integers a and b. By Theorem 3.4.14, we have $FK \supset K^{(aO_K)}$. Moreover, from Lemma 3.4.13, we obtain that

$$2d \geq [FK : \mathbb{Q}] \geq [K^{(aO_K)} : \mathbb{Q}] \geq \frac{h(O_K)\varphi_K(aO_K)}{3}. \tag{3.4.26}$$

Now, by Theorem 3.4.15, we find an extension L/FK such that $(\mathbb{Z}/ab\mathbb{Z})^2 \hookrightarrow E(L)$ and $[L : FK] \leq b$. As above, by Theorem 3.4.14, we have $L \supset K^{(abO_K)}$ and by Lemma 3.4.13, we also get

$$[L : \mathbb{Q}] \geq [K^{(abO_K)} : \mathbb{Q}] \geq \frac{h(O_K)\varphi_K(abO_K)}{3}. \tag{3.4.27}$$

Note that $[L : \mathbb{Q}] = [L : FK][FK : \mathbb{Q}]$. Using Equation (3.4.26), we may write

$$\begin{aligned} d \geq \frac{[FK : \mathbb{Q}]}{2} &= \frac{[L : \mathbb{Q}]}{2[L : FK]} \\ &\geq \frac{[L : \mathbb{Q}]}{2b} \\ &\geq \frac{h(O_K)\varphi_K(abO_K)}{6b} \quad \text{by Equation (3.4.27).} \end{aligned} \tag{3.4.28}$$

Since $(ab)^2 = |abO_K|$, we multiply by $(ab)^2$ each side of (3.4.28) and we get

$$a^2 b \, \frac{h(O_K)\varphi_K(abO_K)}{6} \leq d \, |abO_K|,$$

that is

$$\#E(FK)[tors] = a^2 b \leq \frac{6d}{h(O_K)} \, \frac{|abO_K|}{\varphi_K(abO_K)}. \tag{3.4.29}$$

By Theorem 3.4.16, we have

$$\begin{aligned} \frac{|abO_K|}{\varphi_K(abO_K)} &\ll h(O_K) \log\log |abO_K| \\ &\leq h(O_K) \log\log (a^2 b)^2 \\ &\ll h(O_K) \log\log \#E(FK)[tors]. \end{aligned} \tag{3.4.30}$$

Putting together Equations (3.4.29) and (3.4.30), we have

$$\#E(FK)[tors] \ll d \log\log \#E(FK)[tors], \tag{3.4.31}$$

and therefore

$$\#E(F)[tors] \leq \#E(FK)[tors] \ll d \log\log d. \tag{3.4.32}$$

□

COMMENTS: We have seen that the proof of Lemma 3.4.17 depends on the bound of the zero β and the O-term. In other words, an improvement of this lemma can be done if one can find an explicit bound of β and also have an information on the O-term.

Motivated by this, we provide an upper bound of the Siegel zero in the next chapter. On the other hand, there are no known (unconditional) estimations for the term $O(\frac{t}{\log t})$ which is from the explicit formula (see (3.4.22)). Thus, it is natural to think of this as our future project.

Chapter 4

An explicit bound on Siegel Zeros

We are now ready to prove our main theorem. Let Δ be a negative fundamental discriminant such that $|\Delta| \geq 3 \times 10^8$. Throughout this chapter, we set the following notations:

$$\ell(\Delta) = e^{\frac{\log|\Delta|}{18\,\mathrm{loglog}\,|\Delta|}},$$
$$f(\Delta) = \max\left\{1, \frac{1}{180}\left(\frac{\log|\Delta|}{\mathrm{loglog}\,|\Delta|}\right)^2\right\}.$$

One can notice that $f(\Delta) \leq \ell(\Delta)$ since $e^t - \frac{9}{5}t^2 \geq 0$ for any $t \geq 0$.

4.1 Sum of reciprocal of the norm of ideals

In this section, we set $x = \frac{1}{4}\sqrt{|\Delta|}f(\Delta)$. As in the previous chapter, we use O_K to denote the ring of integers of an imaginary quadratic field K.

4.1.1 Preliminary discussion

Let $\mathfrak{a}$ be an ideal of O_K. Using Proposition 2.3.21, the norm of the ideal $\mathfrak{a} = u[a, \frac{-b+\sqrt{\Delta}}{2}]$ is given by $N(\mathfrak{a}) = u^2a$. Hence, the sum of reciprocal of the norm of an ideal $\mathfrak{a}$ can be written as

$$\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} = \sum_{u^2a\leq x} \frac{\nu(a)}{u^2a}, \tag{4.1.1}$$

where $\nu(a)$ denotes the number of representations of a as $N(\mathfrak{a})$ such that $\mathfrak{a}$ has no rational integer divisors. We have the following result about $\nu(a)$:

Proposition 4.1.1. *The function $\nu(\cdot)$ is multiplicative and satisfies*

$$\nu(p^\alpha) = \begin{cases} 1+\chi_\Delta(p) & \text{if } p \nmid \Delta \text{ or } \alpha = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1.2}$$

Proof. Let $m, n, \alpha \in \mathbb{N}$. Let p be a rational prime and $\mathfrak{a}$ be an ideal of O_K. We first assume that $\alpha = 1$.

1. If $(\frac{\Delta}{p}) = 1$ then by Equation (2.3.10) we have $(N(\mathfrak{a})) = \mathfrak{a}\,\overline{\mathfrak{a}} = (p) = \mathfrak{p}_1\,\mathfrak{p}_2$ (with $\mathfrak{p}_1, \mathfrak{p}_2$ are prime ideals) as (p) splits in O_K. In other words, we get 2 possibilities for an ideal $\mathfrak{a}$ such that $N(\mathfrak{a}) = p$.

2. If $(\frac{\Delta}{p}) = 0$ then by Equation (2.3.10) we get $(N(\mathfrak{a})) = \mathfrak{a}\,\overline{\mathfrak{a}} = (p) = \mathfrak{p}^2$ ($\mathfrak{p}$ prime ideal) as (p) ramifies in O_K. Then we only have 1 possibility for an ideal $\mathfrak{a}$.

3. If $(\frac{\Delta}{p}) = -1$ then $(N(\mathfrak{a})) = \mathfrak{a}\,\overline{\mathfrak{a}} = (p^2)$ since p is inert and (p) is prime ideal in O_K. Then there is no possiblity for an ideal $\mathfrak{a}$ because no rational divisor is allowed.

Next, suppose that $\alpha > 1$ that is p does not divide Δ. In this case, we have $(N(\mathfrak{a})) = (p^\alpha) = (p)\cdots(p)$ (α-times).

1. If $(\frac{\Delta}{p}) = 1$ then $(p) = \mathfrak{p}_1\,\mathfrak{p}_2$ so that $(N(\mathfrak{a})) = \mathfrak{p}_1^\alpha\,\mathfrak{p}_2^\alpha$. That is $\mathfrak{a}$ is either $\mathfrak{p}_1^\alpha$ or $\mathfrak{p}_2^\alpha$ otherwise $\mathfrak{a}$ has a rational divisor. Then we have 2 possibilities for an ideal $\mathfrak{a}$ with norm $N(\mathfrak{a}) = p^\alpha$.

2. If $(\frac{\Delta}{p}) = -1$ then (p) is a prime ideal in O_K. Assume now that α is even. We have $(N(\mathfrak{a})) = \mathfrak{a}\,\overline{\mathfrak{a}} = (p^\alpha) = (p)\cdots\overline{(p)}$ (α-times). Then there is no possibility because $\mathfrak{a}$ is either $(p)^{\frac{\alpha}{2}}$ or $\overline{(p)}^{\frac{\alpha}{2}}$ and so has a rational divisor. In the case that α is odd, there is also no possibility for an ideal $\mathfrak{a}$ because otherwise p^α will be the product of $N((p)) = p^2$ which cannot happen.

Therefore, we have shown that Equation (4.1.2) holds. Now, assume that $N(\mathfrak{a}) = mn$ such that $(m, n) = 1$. Let $\mathfrak{p}$ be a prime ideal factor of $\mathfrak{a}$. We then have $N(\mathfrak{p}) \mid n$ or $N(\mathfrak{p}) \mid m$. Since $(m, n) = 1$, by the unique factorization of the ideal $\mathfrak{a}$, there exist unique ideals $\mathfrak{a}_n, \mathfrak{a}_m$ such that $N(\mathfrak{a}_n) = n$, $N(\mathfrak{a}_m) = m$ and $\mathfrak{a} = \mathfrak{a}_n\,\mathfrak{a}_m$. In fact, if $\mathfrak{a} = \mathfrak{p}_1^{\alpha_1}\cdots\mathfrak{p}_t^{\alpha_t}$, then

$$\mathfrak{a}_n = \prod_{N(\mathfrak{p}_i)\mid n} \mathfrak{p}_i^{\alpha_i} \quad \text{and} \quad \mathfrak{a}_m = \prod_{N(\mathfrak{p}_j)\mid m} \mathfrak{p}_j^{\alpha_j}.$$

Hence, $\nu(mn) = \nu(m)\nu(n)$. □

Remark 7. In terms of $\nu(a)$, the Dedekind zeta function gives us

$$\begin{aligned}\zeta_K(s) &= \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s} = \sum_a \frac{\nu(a)}{(u^2a)^s},\\ &= \sum_{u\geq 1} \frac{1}{u^{2s}} \sum_a \frac{\nu(a)}{a^s},\\ &= \zeta(2s) \prod_p \Big(1 + \frac{\nu(p)}{p^s} + \frac{\nu(p^2)}{p^{2s}} + \cdots\Big),\end{aligned}$$

and the last equality holds because $\nu(a)$ is multiplicative. We then have

$$\begin{aligned}\zeta_K(s) &= \zeta(2s) \prod_p \Big(1 + (1+\chi_\Delta(p))\Big(\frac{p^{-s}}{1-p^{-s}}\Big)\Big),\\ &= \zeta(s) \prod_p \Big(\frac{1}{1+p^{-s}}\Big) \prod_{p,\chi_\Delta(p)=1} \Big(1 + 2\Big(\frac{p^{-s}}{1-p^{-s}}\Big)\Big) \prod_{p,\chi_\Delta(p)=0} \Big(1 + \frac{1}{p^s}\Big),\\ &= \zeta(s) \prod_{p,\chi_\Delta(p)=1} \frac{1}{1-p^{-s}} \prod_{p,\chi_\Delta(p)=-1} \frac{1}{1+p^{-s}},\\ &= \zeta(s) \prod_p \frac{1}{1-\chi_\Delta(p)p^{-s}},\\ &= \zeta(s) L(s,\chi_\Delta),\end{aligned}$$

which is as in Proposition 2.3.23.

Going back to Equation (4.1.1), we now have:

$$\begin{aligned}\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} &= \sum_{u^2a \leq \frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{u^2a} + \sum_{\frac{1}{4}\sqrt{|\Delta|} < u^2a \leq x} \frac{\nu(a)}{u^2a},\\ &\leq \sum_{u\geq 1} \frac{1}{u^2} \sum_{a\leq \frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{a} + \sum_{u\geq 1} \frac{1}{u^2} \sum_{\frac{1}{4}\sqrt{|\Delta|} < a \leq x} \frac{\nu(a)}{a},\\ &\leq \frac{\pi^2}{6}\Big(S_0 + S_1 + S_2\Big),\end{aligned}$$

where

$$S_0 = \sum_{a\leq \frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{a}, \tag{4.1.3}$$

S_1 is the sum of terms $\frac{\nu(a)}{a}$ over $\frac{1}{4}\sqrt{|\Delta|} < a \leq x$ such that a has at least one prime power factor greater than $\ell(\Delta)$, and S_2 is the sum of all the other terms.

In [27, Table 4], Watkins gives all the number of negative fundamental discriminants with class number $N \leq 100$ and the largest absolute value of

such discriminant. Among all these values, $N = 98$ is the class number corresponding to the largest discriminant which is 2 383 747. On the other hand, it is shown in [28] that if $|\Delta| \leq 300\,000\,000$ then the Siegel zero does not exist. Thus, we can assume that $|\Delta| > 300\,000\,000$ and we have

$$h(\Delta) \geq 101. \tag{4.1.4}$$

4.1.2 Estimating the sums

Now, we are going to concentrate on the two sums S_1 and S_2 respectively. For some technical reasons, we will consider two cases on the values of $|\Delta|$. We first assume that $|\Delta| \geq 10^{100}$.

Proposition 4.1.2. *For $|\Delta| \geq 10^{100}$, we have*

$$S_1 \leq 2.648\, S_0. \tag{4.1.5}$$

Before we prove this proposition, we need the following result.

Lemma 4.1.3. *We have:*

1. *for $2 \leq x \leq 3 \times 10^6$,*

$$-0.629 \leq \sum_{p^\alpha \leq x} p^{-\alpha} - \log\log x - \mathcal{C} \leq 10^{-4},$$

2. *for $x > 3 \times 10^6$,*

$$\left| \sum_{p^\alpha \leq x} p^{-\alpha} - \log\log x - \mathcal{C} \right| \leq \frac{0.2}{(\log x)^3} + \frac{1}{\lceil \sqrt{x} \rceil - 1},$$

where $\mathcal{C} = M + \sum_{\alpha \geq 2} \sum_p p^{-\alpha} = 1.03465...$ and M is the Mertens constant.

Proof. The first part of the lemma is obtained by a direct computation. Let us show the second part. We set

$$R(x) = \sum_{p \leq x} \frac{1}{p}.$$

It is shown in [2] that for all $x \geq 2\,278\,383$, we have

$$\left| R(x) - \log\log x - M \right| \leq \frac{0.2}{(\log x)^3}, \tag{4.1.6}$$

where M is the Mertens constant. Notice the following bound: for $x \geq 2$ we have

$$0 \leq \sum_{p^\alpha \leq x} p^{-\alpha} - R(x) \leq \sum_{\alpha \geq 2} \sum_p p^{-\alpha}.$$

The double summation on the right-hand side is convergent. To see this, we can simply write as follows

$$\sum_{\alpha\geq 2}\sum_{p} p^{-\alpha} = \sum_{p}\sum_{\alpha\geq 2} p^{-\alpha} = \sum_{p}\frac{1}{p^2-p}.$$

Now, for the lower bound, we have

$$\sum_{p^\alpha\leq x} p^{-\alpha} - R(x) = \sum_{\alpha\geq 2}\sum_{p} p^{-\alpha} - \sum_{\alpha=2}^{\infty}\sum_{p^\alpha>x} p^{-\alpha},$$

and

$$\sum_{\alpha=2}^{\infty}\sum_{p^\alpha>x} p^{-\alpha} \leq \sum_{p>\sqrt{x}}\frac{1}{p^2-p} \leq \sum_{n>\sqrt{x}}\frac{1}{n^2-n} = \frac{1}{\lceil\sqrt{x}\rceil-1}.$$

Now, by using Equation (4.1.6), we get

$$\left|\sum_{p^\alpha\leq x} p^{-\alpha} - \log\log x - C\right| \leq \frac{0.2}{(\log x)^3} + \frac{1}{\lceil\sqrt{x}\rceil-1},$$

where

$$C = M + \sum_{\alpha\geq 2}\sum_{p} p^{-\alpha}.$$

□

We now show the above proposition about the sum S_1.

Proof of Proposition 4.1.2: From our description of the sum S_1, there exists $p^\alpha > \ell(\Delta)$ such that $a = p^\alpha r$. More precisely, we obtain

$$\begin{aligned}
S_1 &= \sum_{\frac{1}{4}\sqrt{|\Delta|}<p^\alpha r\leq x}\frac{\nu(p^\alpha r)}{p^\alpha r},\\
&\leq \sum_{r<\frac{x}{\ell(\Delta)}}\frac{\nu(r)}{r}\sum_{\frac{1}{4r}\sqrt{|\Delta|}<p^\alpha\leq\frac{x}{r}}\frac{\nu(p^\alpha)}{p^\alpha},\\
&\leq S_0\sum{}''\frac{\nu(p^\alpha)}{p^\alpha},\\
&\leq S_0\sum{}''\frac{2}{p^\alpha}, \quad \text{by Equation (4.1.2)},\\
&\leq 2S_0\sum{}''p^{-\alpha}, \qquad (4.1.7)
\end{aligned}$$

where $\sum''$ is taken over $\delta(\Delta) < p^\alpha \leq Y(\Delta)$ with

$$\delta(\Delta) = \max\left\{\ell(\Delta), \frac{1}{4r}\sqrt{|\Delta|}\right\} \quad \text{and} \quad Y(\Delta) = \frac{1}{4r}\sqrt{|\Delta|}f(\Delta). \qquad (4.1.8)$$

Applying Lemma 4.1.3 to Equation (4.1.8), it follows that

$$\sum_{\delta(\Delta)<p^{\alpha}\leq Y(\Delta)} p^{-\alpha} \leq \log\left(\frac{\log Y(\Delta)}{\log \delta(\Delta)}\right) + 1.280\times 10^{-7} + 0.629, \tag{4.1.9}$$

where

$$\frac{0.2}{\left(\log Y(\Delta)\right)^3} + \frac{1}{\left\lceil\sqrt{Y(\Delta)}\right\rceil - 1} \leq 1.280\times 10^{-7},$$

and the second constant term is from the first part of the lemma for $|\Delta| \geq 10^{100}$. Now, Equation (4.1.9) becomes

$$\begin{aligned}\sum_{\delta(\Delta)<p^{\alpha}\leq Y(\Delta)} p^{-\alpha} &\leq \log\left(\frac{\log f(\Delta) + \log \frac{1}{4r}\sqrt{|\Delta|}}{\log \delta(\Delta)}\right) + 0.630\\ &\leq \log\left(1 + \frac{\log f(\Delta)}{\log \delta(\Delta)}\right) + 0.630\\ &\leq \log 2 + 0.630.\end{aligned}$$

Thus, Equation (4.1.7) becomes

$$S_1 \leq 2.648\, S_0. \tag{4.1.10}$$

□

We now estimate the sum S_2.

Proposition 4.1.4. *For $|\Delta| \geq 10^{100}$, we have*

$$S_2 \leq 1.887\times 10^{-5}. \tag{4.1.11}$$

Proof. Let k_0 be the number of distinct prime factors of a which appears in the sum S_2. Since $\frac{1}{4}\sqrt{|\Delta|} \leq a = \prod_{i=1}^{k_0} p_i^{\alpha_i} \leq \ell(\Delta)^{k_0}$, by the definition of S_2, we get

$$k_0 \geq \frac{\log \frac{1}{4}\sqrt{|\Delta|}}{\log \ell(\Delta)}.$$

That is

$$\begin{aligned}k_0 &\geq \frac{\log \frac{1}{4} + \frac{1}{2}\log |\Delta|}{\frac{\log |\Delta|}{18\log\log |\Delta|}},\\ &= 9\log\log|\Delta| + 18\log\log|\Delta|\,\frac{\log\frac{1}{4}}{\log|\Delta|},\\ &= \log\log|\Delta|\left(9 + 18\,\frac{\log\frac{1}{4}}{\log|\Delta|}\right),\\ &\geq 8.89\log\log|\Delta|, \qquad \text{for } |\Delta| \geq 10^{100}.\end{aligned} \tag{4.1.12}$$

Then, if we write $\sigma = \sum_{p^\alpha < \ell(\Delta)} \nu(p^\alpha) p^{-\alpha}$ and consider the possible permutations of p^α, we have

$$S_2 \leq \sum_{k \geq k_0} \frac{1}{k!}\sigma^k < \frac{1}{k_0!}\sigma^{k_0} e^\sigma, \tag{4.1.13}$$

where the last inequality is obtained by factorizing the term $\frac{\sigma^{k_0}}{k_0!}$. The estimation of σ is given by

$$\begin{aligned}\sigma &= \sum_{p^\alpha \leq \ell(\Delta)} \nu(p^\alpha) p^{-\alpha}, \\ &\leq 2 \sum_{p^\alpha \leq \ell(\Delta)} p^{-\alpha}, \\ &\leq 2\Big(\log\log \ell(\Delta) + 1.035\Big),\end{aligned}$$

where the last inequality comes from Lemma 4.1.3. Moreover, we have

$$\begin{aligned}\log\log \ell(\Delta) &= \log\left(\frac{\log|\Delta|}{18\log\log|\Delta|}\right), \\ &= \log\log|\Delta| - \log\left(18\log\log|\Delta|\right), \\ &\leq \log\log|\Delta| - 4.584, \qquad \text{for } |\Delta| \geq 10^{100}.\end{aligned}$$

Hence,

$$\sigma \leq 2\left(\log\log|\Delta| - 3.549\right) \leq 2\log\log|\Delta|. \tag{4.1.14}$$

Taking the logarithm of Equation 4.1.13, we obtain

$$\log S_2 \leq k_0 \log\sigma + \sigma - \log(k_0!).$$

However, from Stirling's formula, we have $k_0! > \left(\frac{k_0}{e}\right)^{k_0}$. So we may write

$$\begin{aligned}\log S_2 &\leq k_0 \log\sigma + \sigma - \Big(-k_0 + k_0 \log k_0\Big), \\ &\leq k_0\left(\log\frac{\sigma}{k_0} + 1 + \frac{\sigma}{k_0}\right).\end{aligned}$$

From (4.1.12) and (4.1.14), we get $\frac{\sigma}{k_0} \leq \frac{2\log\log|\Delta|}{8.89\log\log|\Delta|} \leq 0.225$. Therefore,

$$\begin{aligned}\log S_2 &\leq k_0(-0.266) \leq -2.364\log\log|\Delta|, \\ &\leq -2\log\log|\Delta|,\end{aligned}$$

that is

$$S_2 \leq (\log|\Delta|)^{-2} \leq 1.887 \times 10^{-5}. \tag{4.1.15}$$

□

Thus, we can establish the following result.

Proposition 4.1.5. *For $|\Delta| \geq 10^{100}$, we have*

$$\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} \leq 6.003\, S_0. \tag{4.1.16}$$

Proof. Using Proposition 4.1.2 and Proposition 4.1.4, we obtain that

$$\begin{aligned}\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} &\leq \frac{\pi^2}{6}\Big(S_0 + S_1 + S_2\Big),\\ &\leq \frac{\pi^2}{6}\Big(S_0 + 2.648 S_0 + 1.887\times 10^{-5}\Big)\\ &\leq \frac{\pi^2}{6}\times 3.649 \times S_0,\\ &\leq 6.003\, S_0.\end{aligned}$$

□

Next, let us consider the case where $3\times 10^8 \leq |\Delta| < 10^{100}$. Note that our previous definition of the sums S_1 and S_2 still holds but now with the new condition on $|\Delta|$. We first evaluate the sum S_1.

Proposition 4.1.6. *For $3\times 10^8 \leq |\Delta| < 10^{100}$, we have*

$$S_1 \leq 2.648\, S_0. \tag{4.1.17}$$

Proof. By defintion, we get that

$$\begin{aligned}S_1 &= \sum_{\frac{1}{4}\sqrt{|\Delta|}<p^\alpha r\leq x} \frac{\nu(p^\alpha r)}{p^\alpha r},\\ &\leq \sum_{r<\frac{x}{\ell(\Delta)}} \frac{\nu(r)}{r} \sum_{\frac{1}{4r}\sqrt{|\Delta|}<p^\alpha\leq \frac{x}{r}} \frac{\nu(p^\alpha)}{p^\alpha},\\ &\leq 2\, S_0 \sum{}'' p^{-\alpha}, \quad \text{by Equation (4.1.2)},\end{aligned}$$

where $\sum''$ is defined in Equation (4.1.8). As in Equation (4.1.9), we have

$$\sum_{\delta(\Delta)<p^\alpha\leq Y(\Delta)} p^{-\alpha} \leq \log\left(1 + \frac{\log f(\Delta)}{\log \ell(\Delta)}\right) + 0.630 \tag{4.1.18}$$

if $\ell(\Delta) \geq 2$. But if $\ell(\Delta) < 2$, then $|\Delta| < 10^{22}$, so $f(\Delta) = 1$ and $Y(\Delta) \leq 2.5\times 10^{10}$. That is

$$S_1 = \sum_{\frac{1}{4}\sqrt{|\Delta|}<p^\alpha r\leq \frac{1}{4}\sqrt{|\Delta|} f(\Delta)} \frac{\nu(p^\alpha r)}{p^\alpha r} = 0.$$

Therefore, it follows from Equation (4.1.18) that

$$S_1 \leq 2\,\Big(\log 2 + 0.630\Big) S_0 \leq 2.648\, S_0. \tag{4.1.19}$$

□

Before estimating the sum S_2, notice that we now have $\ell(\Delta) < 10.5$ since $|\Delta| < 10^{100}$.

Proposition 4.1.7. *For* $3 \times 10^8 \leq |\Delta| < 10^{100}$, *we have*

$$S_2 \leq 1.301 \times 10^{-8}. \tag{4.1.20}$$

Proof. First, observe that the condition on the constant k_0 is given by

$$k_0 \geq \log\log|\Delta|\left(9 + 18\,\frac{\log\frac{1}{4}}{\log|\Delta|}\right) \geq 22.943, \tag{4.1.21}$$

because $|\Delta| \geq 3 \times 10^8$. Moreover, we have that

$$\begin{aligned}\sigma &= \sum_{p^\alpha \leq \ell(\Delta)} \nu(p^\alpha)p^{-\alpha},\\ &\leq 2\sum_{p^\alpha \leq 10} p^{-\alpha},\\ &\leq 3.325.\end{aligned}$$

Now, Equation (4.1.13) gives us

$$\begin{aligned}S_2 &\leq \frac{1}{k_0!}\sigma^{k_0}e^{\sigma},\\ &< \left(\frac{e}{k_0}\right)^{k_0}\sigma^{k_0}e^{\sigma}, \quad \text{by Stirling's formula},\\ &\leq 1.301 \times 10^{-8},\end{aligned}$$

by taking $k_0 = 23$. □

Finally, we get the next result.

Proposition 4.1.8. *For* $3 \times 10^8 \leq |\Delta| < 10^{100}$, *we have*

$$\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} \leq 6.003\ S_0. \tag{4.1.22}$$

Proof. By Proposition 4.1.6 and Proposition 4.1.7, we obtain that

$$\begin{aligned}\sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} &\leq \frac{\pi^2}{6}\big(S_0 + S_1 + S_2\big),\\ &\leq \frac{\pi^2}{6}\big[S_0 + 2.648S_0 + 1.301 \times 10^{-8}\big],\\ &\leq \frac{\pi^2}{6} \times 3.649 \times S_0,\\ &\leq 6.003\ S_0.\end{aligned}$$

□

4.2 Analytic Estimate of integrals

In this section, we apply analytic methods to find a lower estimate of the sum in Equation (4.1.1).

4.2.1 The integral $\mathcal{I}$

Assume from now on that there exists $\beta > 0$ such that $L(\beta, \chi_\Delta) = 0$. For any $x > 0$, we define the integral

$$\mathcal{I} = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \zeta(s+\beta) L(s+\beta, \chi_\Delta) \frac{x^s}{s(s+2)(s+3)} ds. \tag{4.2.1}$$

We have the following result:

Lemma 4.2.1. *We have*

$$\mathcal{I} \leq \frac{x^{1-\beta}}{6} \sum_{N(\mathfrak{a}) \leq x} \frac{1}{N(\mathfrak{a})}. \tag{4.2.2}$$

Before showing this lemma, we first recall the Perron's Formula. To simplify the notation, if $c > 0$ then by $\int_{c-i\infty}^{c+i\infty}$ we mean $\lim_{T\to\infty} \int_{c-iT}^{c+iT}$.

Lemma 4.2.2. *[1, p.243] If y is any positive real number and $c > 0$, we have*

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s} ds = \begin{cases} 1 & \text{if } y > 1, \\ \frac{1}{2} & \text{if } y = 1, \\ 0 & \text{if } 0 < y < 1. \end{cases}$$

Using this lemma, we get the next result.

Lemma 4.2.3. *For any real positive y, we have*

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{y^s}{s(s+2)(s+3)} ds = \begin{cases} 0 & \text{if } 0 < y < 1. \\ \frac{1}{6} - \frac{y^{-2}}{2} + \frac{y^{-3}}{3} & \text{if } y \geq 1. \end{cases} \tag{4.2.3}$$

Proof. By Lemma 4.2.2 with some change of variable, we obtain

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{y^s}{s+2} ds = \begin{cases} y^{-2} & \text{if } y > 1, \\ \frac{1}{2} & \text{if } y = 1, \\ 0 & \text{if } 0 < y < 1. \end{cases} \tag{4.2.4}$$

Similarly, we also have

$$\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{y^s}{s+3}ds=\begin{cases}y^{-3} & \text{if } y>1,\\ \frac{1}{2} & \textit{if } y=1,\\ 0 & \text{if } 0<y<1.\end{cases} \tag{4.2.5}$$

Now, note that

$$\frac{1}{s(s+2)(s+3)}=\frac{1}{6s}-\frac{1}{2(s+2)}+\frac{1}{3(s+3)}. \tag{4.2.6}$$

Hence, combining (4.2.4), (4.2.5) and (4.2.6), we get

$$\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{y^s}{s(s+2)(s+3)}=\begin{cases}0 & \text{if } 0<y<1,\\ \frac{1}{6}-\frac{y^{-2}}{2}+\frac{y^{-3}}{3} & \text{if } y\geq 1.\end{cases}$$

□

We now show the above lemma on the integral $\mathcal{I}$.

Proof of Lemma 4.2.1: Using Proposition 2.3.23, Equation (4.2.1) becomes

$$\begin{aligned}\mathcal{I}&=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\sum_{\mathfrak{a}}\frac{1}{N(\mathfrak{a})^{s+\beta}}\,\frac{x^s}{s(s+2)(s+3)}\,ds,\\ &=\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\sum_{\mathfrak{a}}\frac{1}{N(\mathfrak{a})^{\beta}}\left(\frac{x}{N(\mathfrak{a})}\right)^s\frac{1}{s(s+2)(s+3)}\,ds.\end{aligned}$$

Applying Lemma 4.2.3 with $y=\frac{x}{N(\mathfrak{a})}$, we may write

$$\mathcal{I}=\sum_{N(\mathfrak{a})\leq x}\frac{1}{N(\mathfrak{a})^{\beta}}\left[\frac{1}{6}-\frac{N(\mathfrak{a})^2}{2x^2}+\frac{N(\mathfrak{a})^3}{3x^3}\right]. \tag{4.2.7}$$

Since $\frac{1}{6}-\frac{1}{2y^2}+\frac{1}{3y^3}\leq\frac{1}{6}$ for any $y\geq 1$, we have

$$\begin{aligned}\mathcal{I}&\leq\frac{1}{6}\sum_{N(\mathfrak{a})\leq x}\frac{1}{N(\mathfrak{a})^{\beta}},\\ &\leq\frac{x^{1-\beta}}{6}\sum_{N(\mathfrak{a})\leq x}\frac{1}{N(\mathfrak{a})}.\end{aligned}$$

□

The next result is needed in order to find the lower bound of $\mathcal{I}$. It concerns the growth of the functions $\zeta(s)$ and $L(s,\chi_\Delta)$ on the imaginary axis. We first recall the functional equations of this two functions. These can be found in standard analytic number theory textbooks, for example see [1, p.259 & p.263].

Theorem 4.2.4. *1. For all s we have*

$$\zeta(s) = 2(2\pi)^{s-1}\Gamma(1-s)\sin(\frac{\pi s}{2})\zeta(1-s). \tag{4.2.8}$$

2. Let χ be a primitive character modulo k. We have

$$\Lambda(1-s,\overline{\chi}) = \frac{i^a k^{\frac{1}{2}}}{\tau(\chi)}\Lambda(s,\chi), \tag{4.2.9}$$

such that

$$\Lambda(s,\chi) = \left(\frac{\pi}{k}\right)^{-\frac{(s+a)}{2}}\Gamma\Big(\frac{s+a}{2}\Big)L(s,\chi),$$

where $\tau(\chi) = \sum_{n=1}^{k}\chi(n)\exp(\frac{2i\pi n}{k})$ is the Gauss sums, $a = 0$ if $\chi(-1) = 1$ and $a = 1$ if $\chi(-1) = -1$.

Now, we have the following lemma which will be used later.

Lemma 4.2.5. *1. For any real t with $|t| \geq 3$, we have*

$$|\zeta(it)| \leq \frac{3}{4\pi}\log(|t|)\left|\sinh\Big(\frac{t\pi}{2}\Big)\Gamma(1-it)\right|. \tag{4.2.10}$$

2. For any real t, we have

$$|L(it,\chi)| \leq \sqrt{\frac{|\Delta|}{\pi}}\left|\frac{\Gamma\Big(1-\frac{it}{2}\Big)}{\Gamma\Big(\frac{1+it}{2}\Big)}\right|\Big(\log|\Delta| + \log\Big(e(|t|+14/5)\Big)\Big). \tag{4.2.11}$$

Proof. Using the equation (4.2.8) of Theorem 4.2.4 with $s = it$, we have

$$|\zeta(it)| = \pi^{-1}\left|\sin\Big(\frac{it\pi}{2}\Big)\Gamma(1-it)\zeta(1-it)\right|.$$

Since $\left|\sin\Big(\frac{it\pi}{2}\Big)\right| = \left|i\sinh\Big(\frac{t\pi}{2}\Big)\right|$, we get

$$|\zeta(it)| = \frac{1}{\pi}\left|\sinh\Big(\frac{t\pi}{2}\Big)\right| \times \left|\Gamma(1-it)\zeta(1-it)\right|.$$

From [25, Theorem 1], we have that for $|t| \geq 3$

$$\begin{aligned}|\zeta(it)| &\leq \frac{1}{\pi}\left|\sinh\Big(\frac{t\pi}{2}\Big)\Gamma(1-it)\right| \times \frac{3}{4}\log(|t|),\\ &\leq \frac{3}{4\pi}\log(|t|)\left|\sinh\Big(\frac{t\pi}{2}\Big)\Gamma(1-it)\right|.\end{aligned}$$

Analogously, for the second part of the lemma, we also use the functional equation for the Dirichlet L-functions associated to the character $\chi = \chi_\Delta$.

Note that χ_Δ is a real primitive character modulo $|\Delta|$ by Theorem 2.2.9. Now using Equation (4.2.9) at $s = it$ with $a = 1$, we have

$$\begin{aligned} L(it,\chi) &= \frac{\tau(\chi)}{i\,|\Delta|^{\frac{1}{2}}}\left(\frac{\pi}{|\Delta|}\right)^{-\frac{(2-it)}{2}}\Gamma\Big(\frac{2-it}{2}\Big)L(1-it,\chi)\frac{1}{\Gamma\Big(\frac{1+it}{2}\Big)}\left(\frac{\pi}{|\Delta|}\right)^{\frac{(1+it)}{2}}, \\ &= \frac{\tau(\chi)}{i\,|\Delta|^{\frac{1}{2}}}\left(\frac{\pi}{|\Delta|}\right)^{-1}\left(\frac{\pi}{|\Delta|}\right)^{\frac{1}{2}}L(1-it,\chi)\,\frac{\Gamma\Big(1-\frac{it}{2}\Big)}{\Gamma\Big(\frac{1+it}{2}\Big)}. \end{aligned}$$

Using the fact that $|\tau(\chi)| = \sqrt{|\Delta|}$ and from [10, Theorem 1.2], we then deduce that for any real t

$$|L(it,\chi)| \leq \sqrt{\frac{|\Delta|}{\pi}}\left|\frac{\Gamma\Big(1-\frac{it}{2}\Big)}{\Gamma\Big(\frac{1+it}{2}\Big)}\right|\Big(\log|\Delta| + \log\Big(e(|t|+14/5\Big)\Big). \tag{4.2.12}$$

□

4.2.2 A lower bound of $\mathcal{I}$

Now, in the opposite direction, we aim to find a lower bound of the integral $\mathcal{I}$. By shifting the line of integration to $\mathrm{Re}(s) = -\beta$, where β is the Siegel zero, and using the Residue Theorem, we have

$$\mathcal{I} = \frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} + \frac{1}{2\pi i}\int_{-\beta-i\infty}^{-\beta+i\infty}\zeta(s+\beta)L(s+\beta,\chi_\Delta)\frac{x^s}{s(s+2)(s+3)}ds. \tag{4.2.13}$$

We are allowed to shift the path of integration because the integral converges uniformly in $\mathrm{Re}(s)$ on any fixed vertical strip. So, as we can see our goal is now to estimate the integral on the right hand side. Let us denote it by $\mathcal{J}$.

Proposition 4.2.6. *We have*

$$|\mathcal{J}| \leq 0.104\,x^{-\beta}\sqrt{|\Delta|}\log|\Delta|. \tag{4.2.14}$$

Proof. Setting $s = -\beta + it$, it follows that

$$\begin{aligned} |\mathcal{J}| &= \frac{1}{2\pi}\left|\int_{-\infty}^{+\infty}\zeta(it)L(it,\chi_\Delta)\,\frac{x^{-\beta+it}}{(-\beta+it)(2-\beta+it)(3-\beta+it)}\,dt\right| \\ &\leq \frac{x^{-\beta}}{2\pi}\int_{-\infty}^{+\infty}|\zeta(it)|\,|L(it,\chi_\Delta)|\,\frac{dt}{|-\beta+it|\,|2-\beta+it|\,|3-\beta+it|}dt. \end{aligned}$$

Let us focus on the product $|-\beta+it|\,|2-\beta+it|\,|3-\beta+it|$. By Theorem 1.2.4, we have $\beta \geq 1 - \frac{1.011}{\log|\Delta|} \geq 0.948$ for $|\Delta| \geq 3\times 10^8$. So we assume for

simplicity that $\frac{1}{\sqrt{2}} < \beta < 1$. We then have

$$|-\beta+it| = \sqrt{\beta^2+t^2} \geq \sqrt{\frac{1}{2}+t^2},$$
$$|2-\beta+it| \geq \sqrt{1+t^2} \quad \text{since } 2-\beta \geq 1,$$
$$|3-\beta+it| \geq \sqrt{1+t^2} \quad \text{since } 3-\beta \geq 1.$$

Hence,

$$|\mathcal{J}| \leq \frac{x^{-\beta}}{2\pi}\int_{-\infty}^{+\infty} |\zeta(it)|\,|L(it,\chi_\Delta)|\,\frac{dt}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}.$$

From Equations (4.2.10) and (4.2.11) of Lemma 4.2.5, we have the following.

$$\begin{aligned}|\mathcal{J}| &\leq \frac{x^{-\beta}}{2\pi}\left[\int\limits_{|t|\leq 3} \frac{|\zeta(it)|\,|L(it,\chi_\Delta)|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\,dt + \int\limits_{|t|>3} \frac{|\zeta(it)|\,|L(it,\chi_\Delta)|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\,dt\right]\\
&\leq \frac{x^{-\beta}}{2\pi}\left[\sqrt{\frac{|\Delta|}{\pi}}\int\limits_{|t|\leq 3} \frac{|\zeta(it)|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|\left(\log|\Delta| + \log\left(e(|t|+14/5)\right)\right)dt\right.\\
&\quad \left. + \frac{3\sqrt{|\Delta|}}{4\pi^{3/2}}\int\limits_{|t|>3} \frac{\log(|t|)\left|\sinh\left(\frac{t\pi}{2}\right)\right|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma(1-it)\,\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|\left(\log|\Delta| + \log\left(e(|t|+14/5)\right)\right)dt\right]\end{aligned}$$

Now, we factorise the term $\sqrt{\frac{|\Delta|}{\pi}}$ and after we split the integrals with respect to the term $\log|\Delta| + \log\left(e(|t|+14/5)\right)$. Then,

$$|\mathcal{J}| \leq \frac{x^{-\beta}\sqrt{|\Delta|}}{2\pi^{3/2}}\left[I_1\log|\Delta| + I_2 + \frac{3\log|\Delta|}{4\pi}\,I_3 + \frac{3}{4\pi}\,I_4\right],$$

where, by using Mathematica,

$$I_1 = \int\limits_{|t|\leq 3} \frac{|\zeta(it)|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|dt \leq 0.727,$$

$$I_2 = \int\limits_{|t|\leq 3} \frac{|\zeta(it)|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|\left(\log\left(e(|t|+14/5)\right)\right)dt \leq 1.632,$$

$$I_3 = \int\limits_{|t|>3} \frac{\log(|t|)\left|\sinh\left(\frac{t\pi}{2}\right)\right|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma(1-it)\,\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|dt \leq 1.204,$$

$$I_4 = \int\limits_{|t|>3} \frac{\log(|t|)\left|\sinh\left(\frac{t\pi}{2}\right)\right|}{(1+t^2)\sqrt{\frac{1}{2}+t^2}}\left|\frac{\Gamma(1-it)\,\Gamma\left(1-\frac{it}{2}\right)}{\Gamma\left(\frac{1+it}{2}\right)}\right|\left(\log\left(e(|t|+14/5)\right)\right)dt \leq 4.676.$$

Hence,

$$\begin{aligned}|\mathcal{J}| &\leq \frac{x^{-\beta}\sqrt{|\Delta|}}{2\pi^{3/2}}\left[0.727\log|\Delta| + 1.632 + 1.204\,\frac{3\log|\Delta|}{4\pi} + 4.676\,\frac{3}{4\pi}\right],\\ &\leq \frac{x^{-\beta}\sqrt{|\Delta|}}{2\pi^{3/2}}\left[1.008\log|\Delta| + 2.749\right],\\ &\leq 0.104\,x^{-\beta}\sqrt{|\Delta|}\log|\Delta|, \qquad \text{for } |\Delta| \geq 3\times 10^8.\end{aligned}$$

□

Now, we have the following proposition on the lower bound of the integral $\mathcal{I}$.

Proposition 4.2.7. *We have*

$$\mathcal{I} \geq \frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} - 0.104\,x^{-\beta}\sqrt{|\Delta|}\log|\Delta|. \tag{4.2.15}$$

Proof. By Triangle inequality, we have

$$\begin{aligned}|\mathcal{I}| &= \left|\frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} + \mathcal{J}\right|,\\ &\geq \left|\frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)}\right| - |\mathcal{J}|,\end{aligned}$$

where $|\mathcal{J}|$ is given in Proposition 4.2.6. From Equation (4.2.7), note that the integral $\mathcal{I}$ has a real value. Therefore, we deduce that

$$\mathcal{I} \geq \frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} - 0.104\,x^{-\beta}\sqrt{|\Delta|}\log|\Delta|. \tag{4.2.16}$$

□

4.3 Deduction of the bound

In Section 4.1, we recall that from Proposition 4.1.5 and Proposition 4.1.8, we obtain

$$\sum_{N(\mathfrak{a})\leq x}\frac{1}{N(\mathfrak{a})} \leq 6.003\;S_0 \tag{4.3.1}$$

when $|\Delta| \geq 10^{100}$ and $3\times 10^8 \leq |\Delta| < 10^{100}$. Previously in Section 4.2, we get from Proposition (4.2.7) that

$$\mathcal{I} \geq \frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} - 0.104\,x^{-\beta}\sqrt{|\Delta|}\log|\Delta|. \tag{4.3.2}$$

However, we also see from Proposition 4.2.1 that

$$\mathcal{I} \leq \frac{1}{6}\, x^{1-\beta} \sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})}. \tag{4.3.3}$$

In other words, we see that $\mathcal{I}$ is bounded below and above. Now, we combine Equations (4.3.2) and (4.3.3) and we may write

$$\begin{aligned}
&\frac{1}{6}\, x^{1-\beta} \sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} \\
&\geq \frac{L(1,\chi_\Delta)x^{1-\beta}}{(1-\beta)(3-\beta)(4-\beta)} - 0.104\, x^{-\beta}\sqrt{|\Delta|}\log|\Delta|, &(4.3.4)\\
&\geq \frac{x^{1-\beta}}{(1-\beta)}\left[\frac{L(1,\chi_\Delta)}{(3-\beta)(4-\beta)} - (1-\beta)\, 0.104\, \frac{\sqrt{|\Delta|}}{x}\log|\Delta|\right]. &(4.3.5)
\end{aligned}$$

Recall that $x = \frac{1}{4}\sqrt{|\Delta|}f(\Delta)$ and by class number formula (2.4.2), we have

$$\begin{aligned}
&\frac{1}{6}\, x^{1-\beta} \sum_{N(\mathfrak{a})\leq x} \frac{1}{N(\mathfrak{a})} \\
&\geq \frac{x^{1-\beta}}{(1-\beta)}\left[\frac{\pi\, h(\Delta)}{\sqrt{|\Delta|}(3-\beta)(4-\beta)} - (1-\beta)\frac{0.416\log|\Delta|}{f(\Delta)}\right] &(4.3.6)\\
&\geq \frac{x^{1-\beta}}{(1-\beta)}\left[\frac{\pi\, h(\Delta)}{\sqrt{|\Delta|}(3-\beta)(4-\beta)} - (1-\beta)\frac{74.88\,(\mathrm{loglog}\,|\Delta|)^2}{\log|\Delta|}\right] &(4.3.7)\\
&\geq \frac{x^{1-\beta}}{(1-\beta)}\left[\frac{\pi\, h(\Delta)}{\sqrt{|\Delta|}(3-\beta)(4-\beta)} - (1-\beta)\times 33.87\right], \quad \text{for } |\Delta| \geq 3\times 10^8. &(4.3.8)
\end{aligned}$$

That is

$$\frac{1}{6}\sum_{N(\mathfrak{a})\leq x}\frac{1}{N(\mathfrak{a})} \geq \frac{1}{(1-\beta)}\left[\frac{\pi\, h(\Delta)}{\sqrt{|\Delta|}(3-\beta)(4-\beta)} - (1-\beta)\times 33.87\right]. \tag{4.3.9}$$

Plugging Equation (4.3.1) into the above Equation (4.3.9), we then obtain

$$1.001\; S_0 \geq \frac{1}{(1-\beta)}\left[\frac{\pi\, h(\Delta)}{\sqrt{|\Delta|}(3-\beta)(4-\beta)} - (1-\beta)\times 33.87\right]. \tag{4.3.10}$$

By Theorem 1.2.4, we have that

$$(3-\beta)(4-\beta) \geq 6.262, \tag{4.3.11}$$

since $\beta \geq 1 - \frac{1.011}{\log|\Delta|} \geq 0.948$ for $|\Delta| \geq 3\times 10^8$.

4.3.1 Estimate of S_0

Here, we give a result on the estimate of the sum $S_0 = \sum_{a<\frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{a}$ in terms of the class number $h(\Delta)$. Before doing that, it is crucial to note that if $[a, \frac{-b+\sqrt{\Delta}}{2}]$ is an ideal with $-a < b \leq a$ and $a < \frac{1}{4}\sqrt{|\Delta|}$ then the corresponding quadratic form $ax^2 + bxy + cy^2$ is reduced. Indeed, assume by contradiction that $a > c$, then

$$|\Delta| = 4ac - b^2 \leq 4ac < 4a^2,$$

which contradicts with $a < \frac{1}{4}\sqrt{|\Delta|}$. This implies that there at most $h(\Delta)$ of such ideals.

We now state the following.

Lemma 4.3.1. *We have*

$$S_0 \leq \frac{h(\Delta)}{10}, \tag{4.3.12}$$

where S_0 is defined in Equation (4.1.3)*.*

Proof. From the definition of $\nu(a)$, we have

$$\sum_{a<\frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{a} \leq \sum_{a<\frac{1}{4}\sqrt{|\Delta|}} \frac{2^{\omega(a)}}{a}, \tag{4.3.13}$$

where $\omega(n)$ denotes the number of distinct prime divisors of n with $\omega(1) = 0$. That is, the number of ideals with norm a is bounded above by $2^{\omega(a)}$. However, we also have

$$\sum_{n=1}^{34} 2^{\omega(n)} = 101. \tag{4.3.14}$$

From Equation (4.3.14), we can say that there can only be at most 101 ideals with norm less or equal to 34. Since the class number $h(\Delta)$ is at least 101, we deduce that

$$\sum_{a<\frac{1}{4}\sqrt{|\Delta|}} \frac{\nu(a)}{a} \leq \sum_{n=1}^{34} \frac{2^{\omega(n)}}{n} + \frac{h(\Delta) - 101}{35}. \tag{4.3.15}$$

This is because the sum becomes larger if more small number a are represented as norm of ideals. In other words, we mean that there are 101 ideals with norm between 1 and 34 and the norm of the other ideals has to be at least 35. Thus, we have

$$S_0 \leq 9.2 + \frac{h(\Delta) - 101}{35} \leq \frac{h(\Delta)}{10}, \tag{4.3.16}$$

with $h(\Delta) \geq 101$ by (4.1.4). □

We are now ready to prove Theorem 1.2.5.

Proof of Theorem 1.2.5: We assume by contradiction that $1-\beta < \frac{1.151}{\sqrt{|\Delta|}}$. Using Equation (4.3.11), Equation (4.3.10) becomes

$$1-\beta \geq \frac{1}{S_0}\left[\frac{0.5011\, h(\Delta)}{\sqrt{|\Delta|}} - \frac{33.836}{\sqrt{|\Delta|}}\right]. \tag{4.3.17}$$

Using Lemma 4.3.1 and Equation (4.1.4), we have

$$1-\beta \geq \frac{5.011-3.855}{\sqrt{|\Delta|}}. \tag{4.3.18}$$

Then, we get $1-\beta \geq \frac{1.156}{\sqrt{|\Delta|}}$. However, this contradicts our earlier assumption, and therefore,

$$1-\beta \geq \frac{1.151}{\sqrt{|\Delta|}}. \tag{4.3.19}$$

This proves our main theorem. □

Conclusion

In this thesis, we combined the approach of Goldfeld and Schinzel with recent computational results to show the explicit lower bound $1-\beta \geq 1.151\,|\Delta|^{-\frac{1}{2}}$ where β is a Siegel zero of an imaginary quadratic field. This bound is certainly not optimal, it can be improved further by sharpening the methods in this work.

A potential application of the bound above was done in Chapter 3 which is about the size of the torsion subgroup of CM elliptic curves. We are, however, not able to obtain an explicit bound for this result in this thesis. We can use this as a starting point of our future project.

List of References

[1] T.M. Apostol. *Introduction to Analytic Number Theory. Undergraduate Texts in Mathematics.* Springer-Verlag, New York - Heidelberg, 1976.

[2] J. Bayless and P. Kinlaw. Explicit bounds for the sum of reciprocals of pseudoprimes and Carmichael numbers. *Journal of Integers Sequences, Volume 20*, 2017.

[3] F. Breuer. Torsion bounds for elliptic curves and drinfeld modules. *J. Number Theory 130, p. 1241-1250*, 2010.

[4] P.L. Clark and P. Pollack. The truth about torsion in the CM case. *C. R. Acad. Sci. Paris, Volume 353, p. 683-688*, 2015.

[5] P.L. Clark and P. Pollack. The truth about torsion in the CM case, II. *Q. J. Math 68, p. 1313-1333*, 2017.

[6] H. Cohen. *Advanced Topics in Computational Number Theory.* Springer-Verlag New York, 2000.

[7] H. Cohen. *Number Theory, Volume I: Tools and Diophantine Equations.* Springer, 2007.

[8] D. A. Cox. *Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication.* John Wiley & Sons, 2013.

[9] H. Davenport. *Multiplicative Number Theory, Second Edition. Graduate Texts in Mathematics, 74.* Springer-Verlag, New York - Berlin, 1980.

[10] A. W. Dudek. An explicit result for $|L(1 + it, \chi)|$. *Funct. Approx. Comment. Math. 53, p. 23-29*, 2015.

[11] M.A. Bennett G. Martin, K. O'Bryant and A. Rechnitzer. Explicit bounds for primes in arithmetic progressions. arXiv:1802.00085, 2018.

[12] D. Goldfeld. An asymptotic formula relating the Siegel zero and the class number of quadratics fields. *Annali Della Scuola Normale Superiore Di Pisa, 4^e Série, p. 611-615*, 1975.

[13] D.M. Goldfeld and A. Schinzel. On Siegel's zero. *Annali Della Scuola Normale Superiore Di Pisa, 4^e Série, p. 571-583*, 1975.

[14] W. Haneke. Uber die reellen nullstellen der dirichletschen L-reihen. *Acta Arithmetica 22, p. 391-421*, 1973.

[15] M. Hindry. *Arithmetics*. Springer London Dordrecht, New York - Heidelberg, 2011.

[16] M. Hindry and J. Silverman. Sur le nombre de points de torsion rationnels sur une courbe elliptique. *C. R. Acad. Sci. Paris Sér. I Math. 329, no. 2, p. 97-100*, 1999.

[17] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory - Second Edition*. Springer, 1990.

[18] H. Kadiri. Explicit zero-free regions for Dirichlet L-functions. *Mathematika 64(2), p. 445-474*, 2018.

[19] D. Marcus. *Number Fields*. Springer-Verlag, 1977.

[20] T. Morrill and T. Trudgian. An elementary bound on Siegel zeroes. arXiv:1811.12521, 2018.

[21] J. Pintz. Elementary methods in the theory of L-functions II - on the greatest real zero of a real L-function. *Acta Arithmetica 31, p. 273-289*, 1976.

[22] J. H. Silverman. *The Arithmetic of Elliptic Curves*. Springer-Verlag New York, 2009.

[23] J. H. Silverman and J. T. Tate. *Rational Points on Elliptic Curves*. Springer, 2015.

[24] J. Stopple. *A primer of analytic number theory*. Cambridge University Press, New York, 2003.

[25] T. Trudgian. A new upper bound for $|\zeta(1+it)|$. *Bull. Aust. Math. Soc. 89, p. 259-264*, 2014.

[26] L. C. Washington. *Elliptic curves: Number Theory and Cryptography*. Chapmann and Hall/CRC, 2008.

[27] M. Watkins. Class numbers of imaginary quadratic fields. *Mathematics of Computation. Volume 73, Number 246, p. 907-938*, 2004.

[28] M. Watkins. Real zeros of real odd Dirichlet L-functions. *Mathematics of Computation. Volume 73, Number 245, p. 415-423*, 2004.